PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 : C07K 5/00, 14/00, A61K 38/00 | A1

(11) International Publication Number: **WO 98/00440**

(43) International Publication Date: 8 January 1998 (08.01.98)

(21) International Application Number: PCT/US97/11719

(22) International Filing Date: 3 July 1997 (03.07.97)

(30) Priority Data:
08/674,895 3 July 1996 (03.07.96) US

(71) Applicant: THE SCRIPPS RESEARCH INSTITUTE [US/US]; 10550 North Torrey Pines Road, La Jolla, CA 92037 (US).

(72) Inventors: BLEIL, Jeffrey, D.; 7585 Darden Court, San Diego, CA 92126 (US). HEIN, Mich, B.; 1355 Santa Margarita Drive, Fallbrook, CA 92028 (US). BOOKBINDER, Louis, H.; 4075 Crystal Dawn Drive, No. 204, San Diego, CA 92122 (US).

(74) Agents: FITTING, Thomas et al.; The Scripps Research Institute, 10550 North Torrey Pines Road, TPC-8, La Jolla, CA 92037 (US).

(81) Designated States: AU, CA, JP, European patent (AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE).

Published
With international search report.

(54) Title: IMMUNOCONTRACEPTIVE COMPOSITIONS CONTAINING SPERM ANTIGENS, AND METHODS OF USE

(57) Abstract

The invention relates generally to immunocontraceptive vaccine compositions comprising antigenic polypeptides having mammalian sp56 sperm protein antigenic determinants. The antigenic polypeptides can be synthetic polypeptides, fusion proteins and chimeric molecules. The vaccine is preferably administered to feral mammals in bait or feed formulations. Certain embodiments describe the use of species specific sequences of sp56. Also described are immunocontraceptive methods, antibodies immunoreactive with the sp56 antigenic determinants, and methods for screening for sperm fertility.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece			TR	Turkey
BG	Bulgaria	HU	Hungary	ML	Mali	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MN	Mongolia	UA	Ukraine
BR	Brazil	IL	Israel	MR	Mauritania	UG	Uganda
BY	Belarus	IS	Iceland	MW	Malawi	US	United States of America
CA	Canada	IT	Italy	MX	Mexico	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NE	Niger	VN	Viet Nam
CG	Congo	KE	Kenya	NL	Netherlands	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NO	Norway	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	NZ	New Zealand		
CM	Cameroon			PL	Poland		
CN	China	KR	Republic of Korea	PT	Portugal		
CU	Cuba	KZ	Kazakstan	RO	Romania		
CZ	Czech Republic	LC	Saint Lucia	RU	Russian Federation		
DE	Germany	LI	Liechtenstein	SD	Sudan		
DK	Denmark	LK	Sri Lanka	SE	Sweden		
EE	Estonia	LR	Liberia	SG	Singapore		

Immunocontraceptive Compositions Containing Sperm Antigens, and Methods of Use

Technical Field

The present invention relates generally to the discovery of a mammalian sp56 cDNA nucleic acid encoding an immunocontraceptive protein. The invention further describes immunocontraceptive polypeptides that are useful in both screening methods, diagnostic methods and therapeutic methods related to immunocontraception.

Background

Recent interest in the possible development of oral contraceptive vaccines derives from their potential cost-effectiveness and lack of side effects (Alexander and Bialy, Reprod. Fertil. Devel., 6:273-280 (1994)). Orally administered, gamete-specific vaccines elicit mucosal immune attack restricted to the reproductive organs, eliminating one or both gametes (Mestecky and McGhee, Adv. Immunol., 40:153-245 (1987); McGhee et al., J. Clin. Immunol., 9:175-199 (1989); Stern et al., J. Reprod. Immunol., 22:73-85 (1992); Meinertz et al., Amer. J. Reprod. Immunol., 25:158-162 (1991); McGhee et al., Reprod. Fertil. Devel., 6:369-379 (1994)).

The development of oral contraceptive vaccines is based upon the cellular specificity required for sperm-egg interaction. In order to fertilize the egg, mammalian sperm first bind tightly and specifically to the zona pellucida, a thick, extracellular glycoprotein coat surrounding the egg. Binding involves contact between the sperm head plasma membrane and surface of the zona pellucida and is largely species-specific. For example, in the mouse, sperm recognize ZP3, one of three glycoproteins in the zona pellucida. The domain of ZP3 recognized by sperm plasma membrane is a 3,900 dalton O-linked oligosaccharide. Since the amino acid sequence of ZP3 varies little amongst species, and because of its ubiquitous role in sperm-egg recognition, considerable effort has been expended in developing

immunocontraceptives that target ZP3 (Spifano and Dean, Rep. Fertil. Devel., 6:319-330 (1994).

Although this strategy may be theoretically effective, recent studies suggest that ovarian pathology, including accelerated follicular atresia and massive proliferation of follicle cells, can result from orally immunizing female mice with a B-cell epitope derived from ZP3, and that immune response does not necessarily lead to contraception. Ovarian pathology might be expected after immunizing females with ZP3, since immune attack and removal of growing oocytes from follicles could, in the absence of a "stop growing" signal from the oocyte, lead to unrestricted division of follicle cells and formation of tumors.

An alternative strategy to development of an effective immunocontraceptive is to target a sperm-specific antigen. Immune attack of such an antigen in males would render those animals infertile. Antibody attack of the antigen, even if it occurs during spermatogenesis, should not block or affect early meiotic events in seminiferous tubules and, therefore, would not result in unrestricted cell proliferation. Moreover, immunized females would likely be infertile, due to immune attack of sperm in the female reproductive tract. A number of sperm proteins have been used to immunize animals and/or cause infertility, including the sperm hyaluronidase PH-20 antigen (Primakoff et al., Nature, 335:543-546 (1988); Fiszer-Szafarz et al., Acta. Biochimica Polonica, 42:31-33 (1995); Lin et al., J. Cell Biol., 125:157-163 (1994); and Gmachl et al., FEBS Lett., 336:545-548 (1993)), the sperm lactate dehydrogenase isozyme, LDH-C4 (Goldberg et al., Fertil. Steril., 35:214-217 (1981); Freemerman et al., Mol. Reprod. Dev., 34:140-148 (1993)), the intra-acrosomal protein sp-10 (Srinivasan et al., Biology of Reprod., 53:462-471 (1995)), and sp17 (O'Rand et al., J. Reprod. Immunol., 25:89-102 (1993)).

A mouse sperm protein, called sp56, which has many of the properties expected of the protein which mediates sperm recognition of ZP3 was recently identified (Bleil and Wassarman,

Proc. Natl. Acad. Sci., USA, 87:5563-5567 (1990) and Cheng et al., J. Cell Biol., 125:867-878 (1994). sp56 was first discovered in studies designed to identify the mouse sperm's ZP3 recognition protein. Purified ZP3 was modified with photoactivatable crosslinkers and challenged with live sperm or crosslinked to affinity columns and exposed to total, soluble sperm proteins. Each of these two, independent methods identified only one mouse sperm protein, sp56. sp56 is a homomultimeric peripheral membrane protein, with a monomer size of 56,000 daltons, located solely on the sperm head plasma membrane overlying the acrosome, precisely where ZP3 binding sites are located. sp56 is a lectin having specific affinity for ZP3, ZP3's FD oligo, and a-galactose, a terminal sugar required for sperm recognition of the FD-oligo. Approximately 25,000 sp56 molecules are predicted to be present on the sperm surface based on the number of ZP3 binding sites on sperm.

The cDNA encoding sp56 has recently been cloned (Bookbinder et al., Science, 269:86-89 (1995)). The primary sequence of sp56, derived from the isolated cDNA, indicates that it is a member of a superfamily of protein receptors characterized by the presence of a series of 60 amino acid-long homology repeats called Sushi domains. Each Sushi domain contains ten virtually invariant amino acids. Consistent with its membership in the superfamily, sp56 cDNA also encodes a stretch of 44 amino acids near the carboxy-terminus which is unique to the protein. As such, the unique stretch is referred to as a sp56 specific sequence domain (sss), and is believed to define a species-restricted epitope. As expected of a cDNA encoding a peripheral membrane protein, the cDNA sequence encodes a signal peptide, located upstream of the sp56 amino-termius. The sp56 cDNA encodes a 62,000 dalton polypeptide as determined from the amino-terminus to the end of the cDNA open reading frame. The 62,000 dalton protein, called "pro-sp56", is processed to mature sp56 (56,000 daltons) by proteolytic removal of approximately 6,000 daltons at the carboxy-terminus. The molecules is thus referred

to as sp56 for sperm 56,000 daltons. Monoclonal antibodies and molecular probes developed from sp56 cDNA have been used to demonstrate that sp56 expression is restricted solely to spermatogenic cells of the mouse; neither the protein nor its mRNA accumulate in other mouse tissues.

In contrast to the above-noted sperm antigens, however, sp56 is unique in the following respects: 1) sp56 expression is restricted to sperm, and no closely related protein (or isozyme) is expressed in somatic tissue, minimizing the chance of systemic immunity; 2) sp56 resides on the sperm surface and is, therefore, exposed to immune attack in the reproductive tract of both males and females; and 3) sp56 expression occurs after maturation of the immune system, meaning that immunization of both males and females presents a "foreign" antigen.

In view of the foregoing characteristics of sp56, it is therefore desirable to design an immunocontraceptive vaccine composition based on a sp56 antigenic polypeptide.

Brief Description of the Invention

A mammalian sperm protein, designated sp56, has been discovered to be useful as a target in immunocontraceptive vaccines for any mammalian species from rodents to man.

Thus, the invention describes a variety of antigenic polypeptides that comprise sp56 antigenic determinants useful in immunocontraceptive vaccine compositions.

Also described are methods for immunocontraception using the vaccine composition, including oral immunocontraception using bait or feed formulations containing the antigenic polypeptides.

The invention also describes anti-sp56 antibodies and monoclonal antibodies useful for detecting the presence of sp56 antigens in samples. Immunological methods are contemplated using the antibodies and the antigenic polypeptides for detecting sp56 antigenic determinants and anti-sp56 antibodies. These methods are useful for assessing sperm fertility, detection of

sperm sp56 antigens and confirming effectiveness of the present vaccines at inducing immune responses.

Recombinant DNA expression vectors and method for producing recombinant proteins comprising sp56 antigenic determinants are also described.

The present invention has the following advantages:

The use of an immunocontraceptive is reversible insofar as the level of circulating antibody will decline if there is no repeated exposure to the sp56 antigenic determinant. When formulated as bait for pest control of feral animals, the material is non-toxic insofar as it does not contain toxins and therefore is relatively safe when inadvertently consumed by non-targeted animals or humans. In addition, the immunocontraceptive can be formulated to contain species-restricted sp56 sequences such that contraception is only achieved in the target species, and not in other animals or human who inadvertently consume the bait. Other advantages will be readily apparent to one skilled in the art.

Brief Description of the Drawings

Figure 1 shows the amino acid residue sequence of the mouse sp56 cDNA clone 7.1 having GenBank Accession Number U17108. The signal peptide (leader sequence) begins at -32 and ends with a glycine residue (gly) that is two residues before the first Sushi domain indicated by an underline. The remaining six Sushi domains are similarly indicated.

Figure 2 is a graphic illustration of the sp56 7.1 cDNA clone where the base numbers are indicated as "ruler" markings. Ribosome binding site (RBS) and poly A signal along with the experimentally determined poly A site are indicated on the ruler. The encoded polypeptide regions are indicated by arrows. Positions of the Sushi domains 1 through 7, the signal peptide, the sp56 specific sequence (sss) and the basic C-terminal tail of the pro-sp56 polypeptide are also indicated.

Figure 3 schematically illustrates a typical Sushi domain structure where each domain is composed of approximately 60 amino acids, of which 10 discontinuous amino acids (from the amino to the carboxy terminus include C, P, Y, C, G, C, G, W, P and C, in single-letter code) are virtually invariant and are located at specific positions within the sequence. The appromixately 50 remaining amino acids evidence considerable divergence. This general structure has the following features: 1) Two disulfide bonds of C1-C3 and C2-C4 which forms the looped-loop structure; 2) At least one proline near each disulfide bond, providing chain turns necessary for attachment of contiguous Sushi domains; and 3) The presence of G and W on the inner hydrophobic loop.

Figure 4 illustrates the first portions of the step-wise process of preparing a chimeric gene for encoding a chimeric polypeptide molecule comprising a TMV coat protein (cp) with a sp56 polypeptide, in this case one derived from the sp56 specific sequence (sss) domain. The process, including the steps of preparing the chimeric primers, is more completely described in Example 3.

Figure 5 illustrates the final steps of the step-wise process of preparing a chimeric gene for encoding TMV coat protein (cp) with a sp56 polypeptide by cloning the TMV/sss gene into the homologous region of a complete TMV gene. The expression of the chimeric polypeptide molecule is shown with the TMV having an sp56-specific polypeptide (hapten) exposed on the viral coat. The final steps are more completely described in Example 3.

Figure 6 is a schematic plasmid map of pRLCTA-sp56 that is described in Example 3.

Figure 7 is a schematic plasmid map of pRLCTm-sp56 that is described in Example 3.

Figure 8 is a schematic plasmid map of pRLCTB-1sp56 that is described in Example 3.

Figure 9 is a schematic plasmid map of pCTA02-1sp56 that is described in Example 3.

Figure 10 shows the Western blot analysis of antiserum from a CT-B-sss3pp-immunized mouse on mouse sperm proteins from a nonimmunized mouse under nonreducing conditions on SDS-PAGE. The antisera from the CT-B-sss3pp-immunized mouse, labeled as CTB-SSS, and the positive control monoclonal antibody 7H12 specifically reacted with nonimmunized sperm proteins having a molecular weight of 40,000 daltons (40 kDa).

Detailed Description of the Invention

A. Definitions

Amino Acid Residue: An amino acid formed upon chemical digestion (hydrolysis) of a polypeptide at its peptide linkages. The amino acid residues described herein are preferably in the "L" isomeric form. However, residues in the "D" isomeric form can be substituted for any L-amino acid residue, as long as the desired functional property is retained by the polypeptide. NH_2 refers to the free amino group present at the amino terminus of a polypeptide. COOH refers to the free carboxy group present at the carboxy terminus of a polypeptide. In keeping with standard polypeptide nomenclature (described in J. Biol. Chem., 243:3552-59 (1969) and adopted at 37 CFR §1.822(b)(2)), abbreviations for amino acid residues are shown in the following Table of Correspondence:

TABLE OF CORRESPONDENCE

SYMBOL		AMINO ACID
1-Letter	3-Letter	
Y	Tyr	tyrosine
G	Gly	glycine
F	Phe	phenylalanine
M	Met	methionine
A	Ala	alanine
S	Ser	serine
I	Ile	isoleucine
L	Leu	leucine
T	Thr	threonine
V	Val	valine
P	Pro	proline
K	Lys	lysine
H	His	histidine
Q	Gln	glutamine
E	Glu	glutamic acid
Z	Glx	Glu and/or Gln
W	Trp	tryptophan
R	Arg	arginine
D	Asp	aspartic acid
N	Asn	asparagine
B	Asx	Asn and/or Asp
C	Cys	cysteine
X	Xaa	Unknown or other

It should be noted that all amino acid residue sequences represented herein by formulae have a left- to-right orientation in the conventional direction of amino terminus to carboxy terminus. In addition, the phrase "amino acid residue" is broadly defined to include the amino acids listed in the Table of Correspondence and modified and unusual amino acids, such as those listed in 37 CFR 1.822(b)(4), and incorporated herein by reference. Furthermore, it should be noted that a dash at the

beginning or end of an amino acid residue sequence indicates a peptide bond to a further sequence of one or more amino acid residues or a covalent bond to an amino-terminal group such as NH_2 or acetyl or to a carboxy-terminal group such as COOH.

Recombinant DNA (rDNA) molecule: a DNA molecule produced by operatively linking two DNA segments. Thus, a recombinant DNA molecule is a hybrid DNA molecule comprising at least two nucleotide sequences not normally found together in nature. rDNA's not having a common biological origin, i.e., evolutionarily different, are said to be "heterologous".

Vector: a rDNA molecule capable of autonomous replication in a cell and to which a DNA segment, e.g., gene or polynucleotide, can be operatively linked so as to bring about replication of the attached segment. Vectors capable of directing the expression of genes encoding for one or more polypeptides are referred to herein as "expression vectors".

Antibody: The term antibody in its various grammatical forms is used herein to refer to immunoglobulin molecules and immunologically active portions of immunoglobulin molecules, i.e., molecules that contain an antibody combining site or paratope. Exemplary antibody molecules are intact immunoglobulin molecules, substantially intact immunoglobulin molecules and portions of an immunoglobulin molecule, including those portions known in the art as Fab, Fab', $F(ab')_2$ and F(v).

Antibody Combining Site: An antibody combining site is that structural portion of an antibody molecule comprised of a heavy and light chain variable and hypervariable regions that specifically binds (immunoreacts with) an antigen. The term immunoreact in its various forms means specific binding between an antigenic determinant-containing molecule and a molecule containing an antibody combining site such as a whole antibody molecule or a portion thereof.

Monoclonal Antibody: A monoclonal antibody in its various grammatical forms refers to a population of antibody molecules that contain only one species of antibody combining site capable

of immunoreacting with a particular epitope. A monoclonal antibody thus typically displays a single binding affinity for any epitope with which it immunoreacts. A monoclonal antibody may therefore contain an antibody molecule having a plurality of antibody combining sites, each immunospecific for a different epitope, e.g., a bispecific monoclonal antibody. Although historically a monoclonal antibody was produced by immortalization of a clonally pure immunoglobulin secreting cell line, a monoclonally pure population of antibody molecules can also be prepared by the methods of the present invention.

Upstream: In the direction opposite to the direction of DNA transcription, and therefore going from 5' to 3' on the non-coding strand, or 3' to 5' on the mRNA.

Downstream: Further along a DNA sequence in the direction of sequence transcription or read out, that is traveling in a 3'- to 5'-direction along the non-coding strand of the DNA or 5'- to 3'-direction along the RNA transcript.

Reading Frame: Particular sequence of contiguous nucleotide triplets (codons) employed in translation that define the structural protein encoding-portion of a gene, or structural gene. The reading frame depends on the location of the translation initiation codon.

Polypeptide: A linear series of amino acid residues connected to one another by peptide bonds between the alpha-amino group and carboxy group of contiguous amino acid residues.

Protein: A linear series of greater than 50 amino acid residues connected one to the other as in a polypeptide.

Substantially Purified or Isolated: When used in the context of polypeptides or proteins, the terms describe those molecules that have been separated from components that naturally accompany them. Typically, a monomeric protein is substantially pure when at least about 60% to 75% of a sample exhibits a single polypeptide backbone. Minor variants or chemical modifications typically share the same polypeptide sequence. A substantially purified protein will typically comprise over about 85% to 90% of

a protein sample, more usually about 95%, and preferably will be over about 99% pure. Protein or polypeptide purity or homogeneity may be indicated by a number of means well known in the art, such as polyacrylamide gel electrophoresis of a sample, followed by visualization thereof by staining. For certain purposes, high resolution is needed and high performance liquid chromatography (HPLC) or a similar means for purification utilized.

Synthetic Peptide: A chemically produced chain of amino acid residues linked together by peptide bonds that is free of naturally occurring proteins and fragments thereof.

Fusion Protein: A polypeptide produced by recombinant DNA methods in which a first polypeptide domain is operatively linked to a second polypeptide domain by the peptide bond produced through expression of a single open reading frame to express a single "fused" polypeptide.

Chimeric Molecule: A bifunctional molecule formed by connecting two separate molecule through chemical linkage, such as by crosslinking two isolated polypeptides.

B. Immunocontraceptive Vaccine Compositions

The present invention describes the use of mammalian sperm sp56 antigenic determinants as immunogens for the purpose of inducing an immune response in a subject mammal, thereby producing antibodies which immunoreact with the sp56 antigenic determinant(s) on sperm *in vivo* and thereby influence sperm content and fertility in the mammal.

The present compositions are therefore referred to as vaccines because of the intended purpose of inducing an immune response in the subject, and in that respect are intended to contain typical components of a vaccine, depending upon the route of administration, as described further herein.

The present compositions are also referred to as immunocontraceptive because the mechanism of contraception effected by the present invention is based on the presence of an

immune response in the treated mammal which effects the content and quality of sperm, thereby effecting fertility and acting as a contraceptive. In this context, it is to be appreciated that any effect on sperm fertility is deemed a contraceptive property, and absolute fertility is not required for effectiveness, although is typically preferred in most embodiments of the invention.

The invention is based primarily on the discovery of the mammalian sperm protein sp56 as having a major role in binding to the egg during the fertilization process in mammals, and more particularly the discovery that the sp56 protein is well suited as a target for the immune system in effecting an immunocontraceptive response. The invention is also characterized as identifying a variety of useful epitopes (antigenic determinants) within the sp56 protein which provide different, unique and useful antigenic determinants (targets) for a therapeutic immune response.

An antigenic polypeptide is therefore defined herein as a polypeptide which includes within its amino acid residue sequence a sequence of amino acids that define an sp56 antigenic determinant. An sp56 antigenic determinant can be any portion of the sp56 protein as described herein, which, upon immunization, is capable of inducing antibodies immunoreactive with the immunizing portion of the sp56 protein. As is well known in the antibody arts, an antigenic determinant can be any of a variety of materials in terms of chemical constituency and size. In the present context, the determinant can be a polypeptide of a variety of lengths having sp56 sequences, and typically is at least about 7 amino acids in length up to large proteins of 400 to 600 amino acid residues in length. The size of the determinant is not intended to be limiting other than to identify preferred embodiments.

The present invention describes a large variety of sp56 determinants derived from different portions of the sp56 protein which may be used alone or incombination in a composition of the present invention. Use of polypeptide-defined determinants from

these different portions of sp56 provide different advantages apparent to one skilled in the art. For instance, some regions of sp56 contain species specific sequences (sss) which provide for an immune response specific for a particular species. Compositions restricted to these sss sp56 determinants will produce sterility in only the target species and will not be contraceptive in other species. Alternatively, other regions of sp56 contain sequences which are conserved (homologous) in different mammalian species, and therefore would serve as universal immunocontraceptive determinants. Furthermore, the use of combinations of sp56 determinants in a single composition provide for the production of a battery of different immunoreactivities, providing additional features to an immunocontraceptive vaccine composition of the present invention.

An immunocontraceptive vaccine composition can include a variety of materials to facilitate the vaccine's purpose. In addition to an antigenic polypeptide, the composition can include a pharmaceutically acceptable excipient, carriers, immune stimulators, antigenic molecules, adjuvants, stabilizers and the like materials, as is well known. Also, the antigenic polypeptide can be in a variety of forms, including a synthetic polypeptide, a hybrid molecule, a fusion protein, a recombinant protein, a chimeric polypeptide molecule, an isolated protein, and the like forms which contain an sp56 antigenic determinant.

Synthetic polypeptides are described in detail hereafter. A hybrid molecule is a molecule having two domains: a first polypeptide domain comprising the sp56 antigenic determinant, and a second polypeptide domain having a different function, such as carrier, presentation, delivery, immune stimulator or vehicle for synthesis.

The second polypeptide domain, as described herein, can have a variety of purposes. It can be a carrier protein, to add bulk to the composition. It can be bacterial or viral antigenic protein, to add immunogenicity to the vaccine composition. It can be a bacterial or viral structural protein to facilitate

expression and/or delivery into a preselected formulation. For example, the use of Tobbacco Mosaic Virus (TMV) coat protein in a fusion protein facilitates the spread of expressed sp56 antigenic determinants throughout the tissue of a TMV-infected plant, thereby facilitating the production of a feed that contains the active immunocontraceptive ingredient. The second polypeptide can be a protein with activity in stimulating the immune response. Other functions may be included in the second polypeptide domain and will be apparent to one skilled in the art.

Preferred second polypeptide domains include cholera toxin subunit A, cholera toxin subunit B, diphtheria toxin, influenza virus HA protein, murine leukemia virus coat protein, salmonella surface protein, and tobacco mosaic virus coat protein. These proteins are well characterized in the recombinant DNA arts and are readily avialable for use in fusion proteins or in chimeric polypeptide molecules. Exemplary uses are described further herein.

The first and second polypeptide domains of a hybrid molecule can be linked by a variety of means. If produced by recombinant DNA expression of a single polypeptide, then the first and second polypeptide domains are operatively linked by a peptide bond as a sigle fusion protein. The production of fusion proteins by recombinant DNA methods is also described further herein.

If produced as separate polypeptides, the first and second polypeptides are then operatively linked by chemical cross-linking methods as are well known, and the product is referred to herein as a chimeric polypeptide molecule. Methods for chemical cross-linkages between polypeptides is well known in the chemical and protein arts. Exemplary chimeric polypeptide molecules useful in a vaccine composition of the invention are described herein.

Insofar as sp56 can be isolated from sperm, substantially isolated sp56 protein can be used in a vaccine composition of the

present invention. Methods for isolating native sp56 protein from epididymis tissues, sperm or semen are described in the Examples.

An antigenic polypeptide having an sp56 antigenic determinant of the present invention has a variety of uses according to the present invention.

Thus, a polypeptide of this invention is characterized by its ability to immunologically mimic an epitope (antigenic determinant) expressed by sp56. Such a polypeptide is useful herein as a component in an immunocontraceptive vaccine composition, in an inoculum for producing antibodies that immunoreact with native sp56 and as an antigen in immunologic methods. Representative and preferred sp56 polypeptides for use as an immunogen in an inoculum or vaccine composition are shown herein.

As used herein, the phrase "immunologically mimic" in its various grammatical forms refers to the ability of an antigenic polypeptide of this invention to immunoreact with an antibody of the present invention that recognizes an epitope of sp56 as defined herein.

It should be understood that a subject polypeptide need not be identical to the amino acid residue sequence of sp56, so long as it includes the required sequence.

C. sp56 Proteins and Polypeptides

The cDNA encoding the mouse sperm protein sp56, referred to as such based on the mature processed protein as described below, has been cloned and sequenced as described by Bookbinder et al., Science, 269:86-89 (1995), the disclosures of which are hereby incorporated by reference. In view of the comparison of the encoded sp56 amino acid sequence to known sequences, sp56 is considered to be a new member of a superfamily of protein receptors that contain multiple consensus repeats of approximately 60 amino acids in length, termed Sushi domains, that are more completely described below. Thus, sp56 is

sufficiently similar to other proteins that allows generalizations regarding function, and assures that one can identify and isolate the sp56 gene in any mammalian species.

Thus, although variations at both the amino acid and nucleotide sequence level may exist in mammalian isolates of sp56, such variations are not to be construed as limiting. For example, allelic variation within a mammalian species can tolerate a several percent difference between isolates of a type of sp56, the differences of which comprise non-deleterious variant amino acid residues. Thus a protein of about 95% homology, and preferably at least 98% homology, to a disclosed sp56 protein is considered to be an allelic variant of the disclosed sp56 protein, and therefore is considered to be a sp56 protein of this invention.

The complete coding nucleotide sequence of mouse pre-prosp56 cDNA, referred to as clone 7.1, is 2026 nucleotides in length as listed in SEQ ID NO 1 and in GenBank with Accession Number U17108. The complete pre-prosp56 encoding cDNA clone presents a 1737 nucleotide open reading frame (ORF) with a 32-amino acid N-terminal signal peptide whose cleavage site between amino acid positions -1 and +1 corresponding to a cleavage site after nucleotide position 175 (SEQ ID NO 1).

Translation of this mouse cDNA sequence encodes that an unprocessed pro-prosp56 protein of 579 amino acid residues, referred to as mouse pre-prosp56, shown in SEQ ID NO 2 from position -32 to +547. The amino acid sequence of mouse pre-prosp56 is also listed in SEQ ID NO 1 with the encoding cDNA nucleotide sequence.

Processing of the pre-prospecies to prosp56 results in a 547 amino acid protein having a molecular weight of 62,000 daltons (62 kDa) based on the experimentally determined amino terminus as described in the Examples. The amino acid residue sequence of mouse prosp56 is shown in Figure 1 and in SEQ ID NO 2 beginning at amino acid position 1 and ending at position 547.

The prosp56 species is further processed by posttranslational truncation of the approximately 6000 dalton (70 amino acids) highly basic carboxy-terminal tail to a mature sp56 protein having a molecular weight of 56 kDa. The posttranslational processing event is thought to occur naturally on motile sperm, or during isolation and treatment of epididymal sperm in nondenaturing detergents, and/or upon the receptor-specific binding to the oocyte ligand, zona pellucida 3 or zp3, as discussed in the Background. The mature sp56 amino acid residue sequence of 477 amino acids is listed in SEQ ID NO 2 beginning with aspartic acid at position 1 and ending with a cysteine residue at position 477. The sequence is also shown in Figure 1.

As previously introduced, comparison of the encoded sp56 amino acid sequence to protein databases indicates that sp56 is a member of a superfamily of protein receptors based on the presence of Sushi domain repeats. The prosp56 protein has considerable homology to proteins of the complement system, such as C4BP, C3b receptor, CR1 and CR2 receptors. Complement proteins, like sp56 of this invention, are known to active exo- or auto-proteolytic activity upon binding to other proteins. Cleavage of the molecules has been shown to cause conformational changes which, in turn, activates subsequence biological events, where in the context of sp56 is the single sperm acrosome-mediated specific interaction with oocyte ligands resulting in fertilization.

As a member of the superfamily, sp56 has seven Sushi domains, six of which are contiguous while the seventh follows an intervening region of 44 amino acids that are unique to mouse sp56. As the encoded intervening region has been determined to be unique to mouse sp56 cDNA, the region has been referred to as sp56 species specific sequence also labeled as "sss". The basic carboxy-terminal domain of 70 amino acids completes the encoded prosp56 protein. The seven Sushi domains are indicated in Figure 1 with underlines amino acid regions. The amino acid residue

positions of the Sushi domains as well as the sp56 species specific sequence are indicated in Table 1 along with the corresponding cDNA region as listed in SEQ ID NO 1 that encodes each region. Table 1 is presented and discussed more completely in Example 2.

Sushi domains are comprised of stretches of approximately 60 amino acids containing ten virtually invariant amino acids. Figure 3 shows a typical Sushi domain, in which 10 discontinuous amino acids (from the amino to the carboxy terminus include C, P, Y, C, G, C, G, W, P and C, in single-letter code) are virtually invariant and are located at specific positions within the sequence. The appromixately 50 remaining amino acids evidence considerable divergence. This general structure has the following features: 1) Two disulfide bonds of C1-C3 and C2-C4 which forms the looped-loop structure; 2) At least one proline near each disulfide bond, providing chain turns necessary for attachment of contiguous Sushi domains; and 3) The presence of G and W on the inner hydrophobic loop.

In the looped-loop structure, both a big and a small loop are present within which a stretch of hydrophobic amino acids is located. The "large loop", composed of a stretch of about 25-30 mostly charged amino acids between C1 and C2 diverges considerably within the superfamily and, because of its charged nature, contributes greatly to the protein receptor function of members of the superfamily (Ichinose et al. J. Biol. Chem., 265:13411-13416 (1990)). The large and small loops are respectively referred to as "ll" and "sl". In addition, Sushi domain 6 is characterized by having an additional cysteine residue, and therefore, an additional loop within the large loop. As such, the large loop is further separated by the cysteine residues into a first loop (Sh6fl) and a middle loope (Sh6ml), the sequences of each of which are listed in Table 1.

Consistent with its membership in the superfamily, the sp56 species specific domain has no detectable homology to any other known protein and contains no cysteine residues, eliminating the

possibility of secondary structure determined by disulfide bonds. The sp56 species specific domain is composed of highly charged amino acids, indicating that it resides on the protein's exterior and is a likely B cell epitope. Computer simulation did not reveal helical or β-sheet secondary structure. Therefore, due to its unique sequence, probable location on the sp56 surface, and lack of secondary structure, this domain of sp56 represents an excellent candidate as a hapten designed to induce a species-specific immune response specific to sp56.

In view of the conserved domains and cleavage sites for generating mature sp56 proteins as described above, similar cleavage patterns, three-dimensional structures and resultant protein species are identifiable in other mammals including humans.

Insofar as the present disclosure identifies sp56 from different mammalian species, the present invention is not to be limited to a sp56 protein derived from one or a few mammalian species. Thus, the invention contemplates a mammalian sp56 protein or homolog, which can be derived, by rDNA or biochemical purification from natural sources, from any of a variety of species including man and other primates, rodents including rabbit, rat and mouse, dog, cat, ungulates including pigs, marsupial, and other mammalian species, without limitation. Thus, the term "homolog" refers to any mammalian sp56 protein or polypeptide having similar three-dimensional structure based on the amino acid residue sequence. In other words, the sp56 species of this invention are homologous molecules in view of the amino acid sequence similarity, the presence of repeating consensus Sushi domains, the presence of a species specific sequence, the overall secondary and tertiary structure of the molecule, and the like physical parameters.

Human and agriculturally relevant animal species are particularly preferred homologs to mouse sp56. Exemplary sp56 species identified herein are mouse, rat, squirrel, rabbits,

groundhogs, prairie dogs, raccoon, deer and related ungulates, kangaroos and human sp56.

As disclosed above, sp56 is produced first in vivo in precursor form, and is then processed into smaller polypeptides having biological activity as described herein. Insofar as these different polypeptide forms are contemplated as useful, the term sp56 protein or polypeptide connotes all species of polypeptide having an amino acid residue sequence derived from the sp56 gene. Thus, a sp56 protein of this invention can be in a variety of forms, depending upon the use therefor, as described herein. Thus as used herein, the phrases "sp56 protein" and "sp56 polypeptide" refers to a sp56 molecule having an amino acid residue sequence that comprises an amino acid residue sequence that corresponds, and preferably is identical, to a portion of a mammalian sp56 protein of this invention.

Thus, the present invention also contemplates a sp56 polypeptide that has an amino acid residue sequence that corresponds to the sequence of the sp56 protein shown in SEQ ID NO 1, and includes an amino acid residue sequence represented by a formula selected from the group consisting of the polypeptides shown in Table 1 as described in Example 2. In this embodiment, the polypeptide is further characterized as having the ability to mimic a sp56 epitope and thereby inhibits sp56 biological activity as described herein and in the Background.

Due to the three dimensional structure of a native folded sp56 molecule, the present invention contemplates that multiple regions of sp56 are involved in sp56 receptor function, which multiple and various regions are defined by the various sp56 polypeptides described above. The ability of the above-described polypeptides to inhibit receptor-ligand binding can readily be measured in in vitro binding assays including immunochemical analyses, sperm-egg binding assays and the like along with in vivo methods including assaying for infertility, by detecting the presence of anti-sp56 antibodies and the like as is more completely discussed in the Examples.

Thus, in another embodiment, the invention contemplates sp56 polypeptide compositions that comprise one or more of the different sp56 polypeptides described above which inhibit sp56 receptor function by facilitating immunocontraception, admixed in combinations to provide simultaneous inhibition of multiple contact sites on the sp56 receptor.

Preferred sp56 proteins and polypeptides for use as antigenic polypeptides in immunocontraceptive vaccine compositions are shown in Table 1 in Example 2. Preferred sp56 proteins include pre-prosp56, prosp56 and mature sp56. Preferred portions of sp56 include each complete Sushi domain and the defined large and small loops therein, including the alternative loops formed in the big loop of Sushi domain 6. Particularly preferred sp56 polypeptides include the species specific sequence and fragments thereof as shown in Table 1.

In view of the variety of forms of a sp56 protein of this invention, sp56 proteins and polypeptides are obtained in a variety of ways. In one embodiment, sp56 can be isolated from a natural tissue, such as testis. Exemplary methods, including biochemical purification and affinity purification, for isolating sp56 are described in Example 1.

Alternatively, a sp56 of this invention can be recombinant protein, that is, produced by recombinant DNA (rDNA) methods as described herein. A recombinant sp56 protein need not necessarily be substantially pure, or even isolated, to be useful in certain embodiments, although recombinant production methods are a preferred means to produce a source for further purification to yield an isolated or substantially pure receptor composition. A recombinant sp56 protein can be present in or on a mammalian cell line or in crude extracts of a mammalian cell line. In other embodiments, a recombinant sp56 protein is produced in or on plants or plant cell lines, for subsequent use in feed or bait delivery means. Preferred expression vector

systems for production of sp56 proteins of this invention in this context are described in Section E.

In one embodiment, a sp56 protein is substantially free of other proteins or peptides, so that the purity of a sp56 reagent and freedom from pharmacologically distinct proteins affords use in the diagnostic and therapeutic methods described herein. The recombinant production methods are ideally suited to produce absolute purity in this regard, although biochemical purification methods from natural sources are also contemplated as previously mentioned. In this regard, a sp56 protein is substantially free from other protein molecules if there are insufficient other molecules such that pharmacological cross-reactivity is not detected in conventional screening assays for ligand binding or biological activity.

In addition to the isolated and recombinant sp56 proteins, the invention contemplates analogs of a sp56 protein. An analog is a man-made variant which exhibits the qualities of a sp56 of this invention in terms of immunological reactivity, ligand binding capacity or the like functional properties of a sp56 protein of this invention. An analog can therefore be a cleavage product of sp56, can be a polypeptide corresponding to a portion of sp56, can be sp56 polypeptide in which a membrane anchor has been removed, and can be a variant sp56 sequence in which some amino acid residues have been altered, to name a few permutations.

A sp56 polypeptide is preferably no more than about 200 amino acid residues in length for reasons of ease of synthesis. Thus, it more preferred that a sp56 polypeptide be no more that about 150 amino acid residues, still more preferably no more than about 65 residues, and most preferably less than 30 amino acid residues in length when synthetic methods of production are used.

A subject sp56 polypeptide includes any analog, fragment or chemical derivative of a polypeptide whose amino acid residue sequence is shown herein so long as the polypeptide is capable of mimicking an epitope of sp56. Therefore, a present polypeptide

can be subject to various changes, substitutions, insertions, and deletions where such changes provide for certain advantages in its use. In this regard, a sp56 polypeptide of this invention corresponds to, rather than is identical to, the sequence of a sp56 protein where one or more changes are made and it retains the ability to induce antibodies that immunoreact with a sp56 polypeptide of this invention.

Thus, the term "analog" includes any polypeptide having an amino acid residue sequence substantially identical to a sequence specifically shown herein in which one or more residues have been conservatively substituted with a functionally similar residue and which displays the ability to induce antibody production as described herein. Examples of conservative substitutions include the substitution of one non-polar (hydrophobic) residue such as isoleucine, valine, leucine or methionine for another, the substitution of one polar (hydrophilic) residue for another such as between arginine and lysine, between glutamine and asparagine, between glycine and serine, the substitution of one basic residue such as lysine, arginine or histidine for another, or the substitution of one acidic residue, such as aspartic acid or glutamic acid for another.

The phrase "conservative substitution" also includes the use of a chemically derivatized residue in place of a non-derivatized residue provided that such polypeptide displays the requisite binding activity.

"Chemical derivative" refers to a subject polypeptide having one or more residues chemically derivatized by reaction of a functional side group. Such derivatized molecules include for example, those molecules in which free amino groups have been derivatized to form amine hydrochlorides, p-toluene sulfonyl groups, carbobenzoxy groups, t-butyloxycarbonyl groups, chloroacetyl groups or formyl groups. Free carboxyl groups may be derivatized to form salts, methyl and ethyl esters or other types of esters or hydrazides. Free hydroxyl groups may be derivatized to form O-acyl or O-alkyl derivatives. The imidazole

nitrogen of histidine may be derivatized to form N-im-benzylhistidine. Also included as chemical derivatives are those peptides which contain one or more naturally occurring amino acid derivatives of the twenty standard amino acids. For examples: 4-hydroxyproline may be substituted for proline; 5-hydroxylysine may be substituted for lysine; 3-methylhistidine may be substituted for histidine; homoserine may be substituted for serine; and ornithine may be substituted for lysine. D-amino acids may also be included in place of one or more L-amino acids. Polypeptides of the present invention also include any polypeptide having one or more additions and/or deletions or residues relative to the sequence of a polypeptide whose sequence is shown herein, so long as the requisite activity is maintained.

The term "fragment" refers to any subject polypeptide having an amino acid residue sequence shorter than that of a polypeptide whose amino acid residue sequence is shown herein.

When a polypeptide of the present invention has a sequence that is not identical to the sequence of a sp56 polypeptide, it is typically because one or more conservative or non-conservative substitutions have been made, usually no more than about 30 number percent, more usually no more than 20 number percent, and preferably no more than 10 number percent of the amino acid residues are substituted. Additional residues may also be added at either terminus for the purpose of providing a "linker" by which the polypeptides of this invention can be conveniently affixed to a label or solid matrix, or carrier. Preferably the linker residues do not form a sp56 epitope, i.e., are not similar is structure to a sp56 protein, as described in Sections B and D.

Any peptide of the present invention may be used in the form of a pharmaceutically acceptable salt. Suitable acids which are capable of forming salts with the peptides of the present invention include inorganic acids such as hydrochloric acid, hydrobromic acid, perchloric acid, nitric acid, thiocyanic acid, sulfuric acid, phosphoric acetic acid, propionic acid, glycolic acid, lactic acid, pyruvic acid, oxalic acid, malonic acid,

succinic acid, maleic acid, fumaric acid, anthranilic acid, cinnamic acid, naphthalene sulfonic acid, sulfanilic acid or the like.

Suitable bases capable of forming salts with the peptides of the present invention include inorganic bases such as sodium hydroxide, ammonium hydroxide, potassium hydroxide and the like; and organic bases such as mono-, di- and tri-alkyl and aryl amines (e.g. triethylamine, diisopropyl amine, methyl amine, dimethyl amine and the like) and optionally substituted ethanolamines (e.g. ethanolamine, diethanolamine and the like).

A sp56 polypeptide of the present invention, also referred to herein as a subject polypeptide, can be synthesized by any of the techniques that are known to those skilled in the polypeptide art, including recombinant DNA techniques as described above. Synthetic chemistry techniques, such as a solid-phase Merrifield-type synthesis, are preferred for reasons of purity, antigenic specificity, freedom from undesired side products, ease of production and the like. An excellent summary of the many techniques available can be found in J.M. Steward and J.D. Young, "Solid Phase Peptide Synthesis", W.H. Freeman Co., San Francisco, 1969; M. Bodanszky, et al., "Peptide Synthesis", John Wiley & Sons, Second Edition, 1976 and J. Meienhofer, "Hormonal Proteins and Peptides", Vol. 2, p. 46, Academic Press (New York), 1983 for solid phase peptide synthesis, and E. Schroder and K. Kubke, "The Peptides", Vol. 1, Academic Press (New York), 1965 for classical solution synthesis, each of which is incorporated herein by reference. Appropriate protective groups usable in such synthesis are described in the above texts and in J.F.W. McOmie, "Protective Groups in Organic Chemistry", Plenum Press, New York, 1973, which is incorporated herein by reference.

In general, the solid-phase synthesis methods contemplated comprise the sequential addition of one or more amino acid residues or suitably protected amino acid residues to a growing peptide chain. Normally, either the amino or carboxyl group of the first amino acid residue is protected by a suitable,

selectively removable protecting group. A different, selectively removable protecting group is utilized for amino acids containing a reactive side group such as lysine.

Using a solid phase synthesis as exemplary, the protected or derivatized amino acid is attached to an inert solid support through its unprotected carboxyl or amino group. The protecting group of the amino or carboxyl group is then selectively removed and the next amino acid in the sequence having the complimentary (amino or carboxyl) group suitably protected is admixed and reacted under conditions suitable for forming the amide linkage with the residue already attached to the solid support. The protecting group of the amino or carboxyl group is then removed from this newly added amino acid residue, and the next amino acid (suitably protected) is then added, and so forth. After all the desired amino acids have been linked in the proper sequence, any remaining terminal and side group protecting groups (and solid support) are removed sequentially or concurrently, to afford the final polypeptide.

A sp56 polypeptide can be used, inter alia, in the diagnostic methods and systems of the present invention to detect a sp56 receptor or sp56 itself present in a body sample, or can be used to prepare an inoculum as described herein for the preparation of antibodies that immunoreact with conserved epitopes on sp56.

In addition, certain of the sp56 polypeptides of this invention can be used in the therapeutic methods of the present invention to inhibit sp56 function as described further herein.

Labels, solid matrices and carriers that can be used with the polypeptides of this invention are described hereinbelow.

Amino acid residue linkers are usually at least one residue and can be 40 or more residues, more often 1 to 10 residues, but do not form a sp56 epitope. Typical amino acid residues used for linking are tyrosine, cysteine, lysine, glutamic and aspartic acid, or the like. In addition, a subject polypeptide can differ, unless otherwise specified, from the natural sequence of

a sp56 protein by the sequence being modified by terminal-NH_2 acylation, e.g., acetylation, or thioglycolic acid amidation, by terminal-carboxlyamidation, e.g., with ammonia, methylamine, and the like.

When coupled to a carrier to form what is known in the art as a carrier-hapten conjugate, a sp56 polypeptide of the present invention is capable of inducing antibodies that immunoreact with sp56. In view of the well established principle of immunologic cross-reactivity, the present invention therefore contemplates antigenically related variants of the polypeptides shown herein.

An "antigenically related variant" is a subject polypeptide that is capable of inducing antibody molecules that immunoreact with a polypeptide described herein and with a sp56 protein of this invention.

D. sp56 Fusion Proteins

In one embodiment, the invention contemplates fusion proteins that contain an antigenic polypeptide of the present invention for use in an immunocontraceptive vaccine composition and for use in other methods where a determinant is useful as disclosed herein.

A fusion protein includes a first polypeptide domain operatively linked to a second polypeptide domain, wherein the first polypeptide domain is an amino acid residue sequence that comprises an antigenic polypeptide of this invention, i.e., one that includes a sequence that defines an sp56 antigenic determinant. The second polypeptide domain can be any polypeptide sequence, but preferably provides some function and advantage to the fusion protein.

Exemplary functions include acting as a carrier, as a presentation structure, as a binding entity for controlling association of the fusion protein with a larger structure, such as an assembled virus particle, as an immune stimulator, such as a protein which augments the immune response, as a antigenic molecule, e.g., to promote a stronger immune response, as an

adjuvant, and the like functions. The invention is not deemed to be limiting as to the function of the second polypeptide domain.

By operatively linked is meant, in the context of fusion proteins, that the two polypeptides are joined, irrespective of order, by the normal polypeptide bond found between two amino acids in a protein, produced by virtue of recombinant DNA (rDNA) expression of a single polypeptide from a gene produced by rDNA methods in which two separate genes have been fused into a single open reading frame. The order in the single polypeptide of the first and second domains need not be sequential relative to the amino- and carboxy- termini of the polypeptide; either the first or second domain can be at the amino-terminus, and vice versa.

Preferred fusion proteins have as a second polypeptide domain a sequence selected from the group consisting of cholera toxin subunit A, cholera toxin subunit B, diphtheria toxin, influenza virus HA protein, murine leukemia virus coat protein, salmonella surface protein, and tobacco mosaic virus coat protein. Exemplary fusion proteins are described in the Examples.

A fusion protein is produced by well known rDNA methods for manipulation and expression of recombinant gene products. Any of a variety of vectors and host systems for expression of a fusion protein may be used, and will determine the route of harvesting and purification of the expressed fusion protein from the host expression system. Preferred are bacterial, yeast and plant expression systems. Plant expression vectors are particularly preferred where the produced fusion protein is to be formulated with a plant material in a bait or feed, and are described further herein.

E. Vectors for Expressing sp56 Proteins and Fusion Proteins

A sp56 protein of this invention can be prepared by a variety of means, although expression in a mammalian cell using a

rDNA expression vector is preferred. Exemplary production methods for a recombinant sp56 are described in the Examples.

Thus, the invention also provides a method for the production of recombinant isolated sp56 proteins, either as intact sp56 protein, as fusion proteins or as smaller polypeptide fragments of sp56. The production method generally involves inducing cells to express a sp56 protein of this invention, recovering the expressed sp56 from the resulting cells, and purifying the expressed sp56 so recovered by biochemical fractionation methods, using a specific antibody of this invention, or other chemical procedures. Inducing expression of a recombinant sp56 protein can comprise inserting a rDNA vector encoding a sp56 protein, or fragment thereof, of this invention, which rDNA is capable of expressing a sp56, into a suitable host cell, and expressing the vector's sp56 gene.

Thus, to facilitate expression of a recombinant sp56 protein or fusion protein of the present invention, DNA segments encoding sp56 or portions thereof are inserted into an expression vector. DNA segments are characterized as including a DNA sequence that encodes a sp56 protein of this invention. That is, the DNA segments of the present invention are characterized by the presence of some or all of a sp56 structural gene. Preferably the gene is present as an uninterrupted linear series of codons where each codon codes for an amino acid residue found in the sp56 protein, i.e., a gene free of introns.

One preferred embodiment is a DNA segment that codes for an amino acid residue sequence that defines a sp56 protein as defined herein, and the DNA segment is capable of expressing a sp56 protein of this invention. A preferred DNA segment codes for an amino acid residue sequence substantially the same as, and preferably consisting essentially of, an amino acid residue sequence or portions thereof shown in SEQ ID NO 2. Particularly preferred DNA segments have a nucleotide sequence derived from the sequence shown in SEQ ID NO 1. Representative and preferred

DNA segments are further described in the Examples and listed in Table 1.

Homologous DNA and RNA sequences that encode the above sp56 protein are also contemplated.

The amino acid residue sequence of a protein or polypeptide is directly related via the genetic code to the deoxyribonucleic acid (DNA) sequence of the structural gene that codes for the protein. Thus, a structural gene or DNA segment can be defined in terms of the amino acid residue sequence, i.e., protein or polypeptide, for which it codes.

An important and well known feature of the genetic code is its redundancy. That is, for most of the amino acids used to make proteins, more than one coding nucleotide triplet (codon) can code for or designate a particular amino acid residue. Therefore, a number of different nucleotide sequences may code for a particular amino acid residue sequence. Such nucleotide sequences are considered functionally equivalent since they can result in the production of the same amino acid residue sequence in all organisms. Occasionally, a methylated variant of a purine or pyrimidine may be incorporated into a given nucleotide sequence. However, such methylations do not affect the coding relationship in any way.

A nucleic acid is any polynucleotide or nucleic acid fragment, whether it be a polyribonucleotide of polydeoxyribonucleotide, i.e., RNA or DNA, or analogs thereof. In preferred embodiments, a nucleic acid molecule is in the form of a segment of duplex DNA, i.e, a DNA segment, although for certain molecular biological methodologies, single-stranded DNA or RNA is preferred.

DNA segments are produced by a number of means including chemical synthesis methods and recombinant approaches, preferably by cloning or by polymerase chain reaction (PCR). DNA segments that encode portions of a sp56 protein can easily be synthesized by chemical techniques, for example, the phosphotriester method of Matteucci, et al., (J. Am. Chem. Soc., 103:3185-3191, 1981) or

using automated synthesis methods. In addition, larger DNA segments can readily be prepared by well known methods, such as synthesis of a group of oligonucleotides that define the DNA segment, followed by hybridization and ligation of oligonucleotides to build the complete segment. Alternative methods include isolation of a preferred DNA segment by PCR with a pair of oligonucleotide primers.

Of course, by chemically synthesizing any desired modifications can be made simply by substituting the appropriate bases for those encoding the native amino acid residue sequence.

Furthermore, DNA segments consisting essentially of structural genes encoding a sp56 protein can be obtained from recombinant DNA molecules containing a gene that defines a sp56 protein of this invention, and can be subsequently modified, as by site-directed or random mutagenesis, to introduce any desired substitutions.

1. Cloning sp56 Genes

A sp56 gene of this invention can be cloned by a variety of cloning methods and from any mammalian species. Obtaining sp56 homologs by cloning is based on the observation that sp56 is a member of a superfamily of protein receptors characterized by the presence of conserved domains. Cloning sp56 genes can be conducted according to the general methods described in the Examples, using nucleic acid homology strategies.

Preferred cloning strategies for isolating a nucleic acid molecule that encodes a sp56 protein of this invention are described in the Examples, and includes the recitation of polynucleotide probes useful for the screening of libraries of nucleic acid molecules believed to contain a target sp56 gene. Particularly preferred probes for isolating a human sp56 gene are provided in Example 6.

Sources of libraries for cloning a sp56 gene of this invention can include genomic DNA or messenger RNA (mRNA) in the

form of a cDNA library from a tissue believed to express a sp56 of this invention. The preferred tissue is testicular tissue.

2. Expression Vectors

In addition, the invention contemplates a recombinant DNA molecule (rDNA) containing a DNA segment of this invention encoding a sp56 protein as described herein. A rDNA can be produced by operatively linking a vector to a DNA segment of the present invention.

As used herein, the term "vector" refers to a DNA molecule capable of autonomous replication in a cell and to which another DNA segment can be operatively linked so as to bring about replication of the attached segment. A vector adapted for expression of a gene product and capable of directing the expression of a sp56 gene is referred to herein as an "expression vector". Thus, a recombinant DNA molecule is a hybrid DNA molecule comprising at least two nucleotide sequences not normally found together in nature.

The choice of vector to which a DNA segment of the present invention is operatively linked depends directly, as is well known in the art, on the functional properties desired, e.g., protein expression, and the host cell to be transformed, these being limitations inherent in the art of constructing recombinant DNA molecules. However, a vector contemplated by the present invention is at least capable of directing the replication, and preferably also expression, of a sp56 structural gene included in DNA segments to which it is operatively linked.

Both prokaryotic and eukaryotic expression vectors are familiar to one of ordinary skill in the art of vector construction and are described by Ausebel, et al., In Current Protocols in Molecular Biology, Wiley and Sons, New York (1993) and by Sambrook et al., Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory, (1989).

In one embodiment, a vector contemplated by the present invention includes a procaryotic replicon, i.e., a DNA sequence

having the ability to direct autonomous replication and maintenance of the recombinant DNA molecule extrachromosomally in a procaryotic host cell, such as a bacterial host cell, transformed therewith. Such replicons are well known in the art. In addition, those embodiments that include a procaryotic replicon also include a gene whose expression confers drug resistance to a bacterial host transformed therewith. Typical bacterial drug resistance genes are those that confer resistance to ampicillin or tetracycline.

Those vectors that include a procaryotic replicon can also include a procaryotic promoter capable of directing the expression (transcription and translation) of a sp56 gene in a bacterial host cell, such as E. coli, transformed therewith. A promoter is an expression control element formed by a DNA sequence that permits binding of RNA polymerase and transcription to occur. Promoter sequences compatible with bacterial hosts are typically provided in plasmid vectors containing convenient restriction sites for insertion of a DNA segment of the present invention. Typical of such vector plasmids are pUC8, pUC9, pBR322 and pBR329 available from Biorad Laboratories, (Richmond, CA), pRSET available from Invitrogen (San Diego, CA) and pPL and pKK223 available from Pharmacia, Piscataway, N.J.

Expression vectors compatible with eucaryotic cells, preferably those compatible with vertebrate cells, can also be used to form the recombinant DNA molecules of the present invention. Eucaryotic cell expression vectors are well known in the art and are available from several commercial sources. Typically, such vectors are provided containing convenient restriction sites for insertion of the desired DNA segment. Typical of such vectors are pSVL and pKSV-10 (Pharmacia), pBPV-1/pML2d (International Biotechnologies, Inc.), pTDT1 (ATCC, #31255), pRc/CMV (Invitrogen, Inc.), the preferred vector pcDNA3 (Invitrogen) described in the Examples, and the like eucaryotic expression vectors.

An alternative expression system that can be used to express a protein of the invention is an insect system. In one such system, Autographa californica nuclear polyhedrosis virus (AcNPV) is used as a vector to express foreign genes. The virus grows in Spodoptera frugiperda cells. The polypeptide-encoding nucleotide sequence may be cloned into non-essential regions (in Spodoptera frugiperda for example the polyhedrin gene) of the virus and placed under control of an AcNPV promoter (for example the polyhedrin promoter). Successful insertion of the polypeptide-encoding nucleotide sequence inactivates the polyhedrin gene and production of non-occluded recombinant virus (i.e., virus lacking the proteinaceous coat coded for by the polyhedrin gene). These recombinant viruses are then used to infect cells in which the inserted gene is expressed. See Smith, et al., J. Biol. Chem., 46:584, (1983); Smith, U.S. Patent No. 4,215,051.

Mammalian cell systems that utilize recombinant viruses or viral elements to direct expression may be engineered. For example, when using adenovirus expression vectors, the coding sequence of a polypeptide may be ligated to an adenovirus transcription/translation control complex, e.g., the late promoter and tripartite leader sequence. This chimeric gene may then be inserted into the adenovirus genome by in vitro or in vivo recombination. Insertion in a non-essential region of the viral genome (e.g., region E1 or E3) will result in a recombinant virus that is viable and capable of expressing the polypeptide in infected hosts (e.g., see Logan & Shenk, Proc. Natl. Acad. Sci., USA, 81:3655-3659, (1984)). Alternatively, the vaccinia virus 7.5K promoter may be used. (e.g., see, Mackett, et al., Proc. Natl. Acad. Sci., USA, 79:7415-7419, (1982); Mackett, et al., J. Virol., 49:857-864, (1984); Panicali, et al., Proc. Natl. Acad. Sci., USA, 79:4927-4931, (1982)). Of particular interest are vectors based on bovine papilloma virus which have the ability to replicate as extrachromosomal elements (Sarver, et al., Mol. Cell. Biol., 1:486, (1981)). Shortly after entry of this DNA into mouse cells, the plasmid replicates to about 100 to 200

copies per cell. Transcription of the inserted cDNA does not require integration of the plasmid into the host's chromosome, thereby yielding a high level of expression. These vectors can be used for stable expression by including a selectable marker in the plasmid, such as the neo gene. Alternatively, the retroviral genome can be modified for use as a vector capable of introducing and directing the expression of the polypeptide-encoding nucleotide sequence in host cells (Cone & Mulligan, Proc. Natl. Acad. Sci., USA, 81:6349-6353, (1984)). High level expression may also be achieved using inducible promoters, including, but not limited to, the metallothionine IIA promoter and heat shock promoters.

For long-term, high-yield production of recombinant proteins, stable expression is preferred. Rather than using expression vectors which contain viral origins of replication, host cells can be transformed with a cDNA controlled by appropriate expression control elements (e.g., promoter and enhancer sequences, transcription terminators, polyadenylation sites, etc.), and a selectable marker. As mentioned above, the selectable marker in the recombinant plasmid confers resistance to the selection and allows cells to stably integrate the plasmid into their chromosomes and grow to form foci which in turn can be cloned and expanded into cell lines.

For example, following the introduction of foreign DNA, engineered cells may be allowed to grow for 1-2 days in an enriched media, and then are switched to a selective media. A number of selection systems may be used, including but not limited to the herpes simplex virus thymidine kinase (Wigler, et al., Cell, 11:223, (1977)), hypoxanthine-guanine phosphoribosyltransferase (Szybalska & Szybalski, Proc. Natl. Acad. Sci., USA, 48:2026, (1962)), and adenine phosphoribosyltransferase (Lowy, et al., Cell, 22:817, (1980)) genes, which can be employed in tk^-, $hgprt^-$ or $aprt^-$ cells respectively. Also, antimetabolite resistance-conferring genes can be used as the basis of selection; for example, the genes for

dhfr, which confers resistance to methotrexate (Wigler, et al., Proc. Natl. Acad. Sci., USA ,77:3567, (1980); O'Hare, et al., Proc. Natl. Acad. Sci., USA, 78:1527, (1981)); gpt, which confers resistance to mycophenolic acid (Mulligan & Berg, Proc. Natl. Acad. Sci., USA, 78:2072, (1981)); neo, which confers resistance to the aminoglycoside G-418 (Colberre-Garapin, et al., J. Mol. Biol., 150:1, (1981)); and hygro, which confers resistance to hygromycin (Santerre, et al., Gene, 30:147, (1984)). Recently, additional selectable genes have been described, namely trpB, which allows cells to utilize indole in place of tryptophan; hisD, which allows cells to utilize histinol in place of histidine (Hartman & Mulligan, Proc. Natl. Acad. Sci., USA, 85:804, (1988)); and ODC (ornithine decarboxylase) which confers resistance to the ornithine decarboxylase inhibitor, 2-(difluoromethyl)-DL-ornithine, DFMO (McConlogue L., In: Current Communications in Molecular Biology, Cold Spring Harbor Laboratory ed., (1987).

In another preferred embodiment, expression vectors compatible for use with plant cells are used to express sp56 in plants. Plants provide advantageous expression and delivery aspects in that a large supply of bulk protein with universal access is readily made from which the protein is either isolatable therefrom or ingestible therein. Thus, transgenic plants containing expression vectors for encoding a recombinant sp56 protein of this invention is useful for preparing sp56 antigenic polypeptides for use in immunocontraceptive vaccine compositions.

Typical expression vectors useful for expression of genes in plants are well known in the art. Typical methods for introducing sp56-encoding genes via expression vectors into plants include Agrobacterium tumefaciens-mediated transformation, plant virus transfection, protoplast transformation, gene transfer into pollen, injection into reproductive organs, injection into immature embryos, and direct insertion, a process referred to as "biolystics". In the case of infection by plant

viruses, a sp56 protein can be produced at high concentrations and isolated at low cost, with the genetic stocks being easily maintained for long periods of time without passaging through plants.

Preferred plants for such expression include any plant for which a compatible expression vector system exists and which can be used to deliver a sp56 immunocontraceptive antigenic polypeptide of this invention to mammals. Exemplary plants include dicots and monocots. Particularly preferred plants include alfalfa, tomato, petunia, soy bean, tobacco, corn, wheat, rice, spinach, asparagus, and the like. Exemplary plant expression vector systems for expression of a recombinant sp56 protein of this invention include those, such as binary vector system utilizing Agrobacterium tumefaciens, described in US Patent Number 5,202,422 and An et al., Plant Molecular Biology Manual, A3:1-19 (1988), the disclosures of which are hereby incorporated by reference.

Other preferred plant vector systems include the pRTL2-series of vectors described by Freed and Carrington, J. Virol., 64: 1590 (1988). These vectors are referred to as shuttle vectors as they are used to generate in-frame recombinant sp56 DNA segments operatively linked to expression control sequences that are then transferred as a single fragment to a plant expression vector that is then introduced into a selected plant. The pRTL2 vectors are derived from pRTL-GUS by removing the GUS coding sequence and replacing it with the coding sequence of interest. As described in the Examples, the GUS sequence has been replaced with the genes encoding cholera toxin α or β chains to provide for the expression of a fusion protein with a sp56 protein of this invention. The pRTL2 vectors have the following elements which are useful as shuttled expression elements to mediate the expression of sp56 in a plant expression vector: 35S promoter derived from 35S transcript gene of cauliflower mosaic virus; tobacco etch virus untranslated leader sequence that serves as translational enhancer; and 35S polyadenylation

sequence of the 3' termius of 35S transcript gene of cauliflower mosaic virus to direct polyadenylation of the transgene transcript. These elements are engineered into a single cassette isolated by HindIII restriction digest. The isolated fragment containing the control elements with the cholera gene system/sp56 coding region are then inserted into a preferred plant expression vector, such as pGA482 (An et al., Plant Molecular Biology Manual, A3:1-19 (1988). The recombinant plant vector is then used to transform a plant virus that then is used to infect plants, such as tobacco or alfalfa. A preferred plant virus is Agrobacterium strain LBA4404.

A further preferred plant delivery expression system relies on a modified tobacco mosaic virus (TMV) to direct the synthesis of TMV coat protein-sp56 chimeric polypeptide molecules on the surface of TMV. The method of using this system to induce an autoantibody response for mediating immunocontraception has been described by Fitchen et al., Vaccine, 13:1051-1057 (1995). The authors described the production of TMV hybrid molecules having oocyte-derived antigens for use as parenterally delivered immunogens. TMV is a preferred antigenic carrier molecule in that it is a self-assembling virus that aggregatres into rod-like particles that then accumulate in virus-infected leaves. The protein component of the assembled particle is the coat protein (cp) into the carboxy-terminal portion of which is inserted non-TMV epitopes. Following infection and expression, the TMV cp and assembled TMV particles provide an antigenic system delivering the antigenic epitope while being immunogenic itself. Coat proteins are readily isolated and presented in particulate or aggregate form. This aspect provides the advantage for effective parenteral or oral immunization of high concentrations of antigen as described in the methods of this invention.

The preparation of exemplary TMV-sp56 chimeric polypeptide molecules of this invention are described in Example 3.

Other preferred plant expression vector systems include pMON530 (Rogers et al., Methods in Enzymol., 153;253 (1987);

pKLYX5,6, and 7 series (Schardl et al., Gene, 61:1 (1987); pBIN19 commercially available from Clontech (Palo Alto, CA).

F. Methods of Immunocontraception

The present invention contemplates immunocontraceptive vaccine compositions useful for inducing fertility in a mammal by inducing an immune response to sperm sp56 protein. The vaccine compositions of the present invention can contain a physiologically tolerable excipient together with an antigenic polypeptide comprising an sp56 antigenic of this invention, dissolved or dispersed therein as an active ingredient.

As used herein, the terms "pharmaceutically acceptable", "physiologically tolerable" and grammatical variations thereof, as they refer to compositions, carriers, diluents and reagents, are used interchangeably and represent that the materials are capable of administration to or upon a mammal without the production of undesirable physiological effects such as nausea, dizziness, gastric upset and the like.

The preparation of a pharmacological composition that contains active ingredients dissolved or dispersed therein is well understood in the art. Typically such compositions are formulated depending on the intended route of administration. As an oral immunocontraceptive, the antigenic polypeptide of this invention is typically included in a bait or feed. In contrast, for targeted (directly administered vaccines) the antigenic polypeptide is prepared as an injectable either as liquid solutions or suspensions, however, solid forms suitable for solution, or suspensions, in liquid prior to use can also be prepared. The preparation can also be emulsified.

The active antigenic polypeptide can be mixed with excipients which are pharmaceutically acceptable and compatible with the active ingredient and in amounts suitable for use in the therapeutic methods described herein. Suitable excipients are, for example, water, saline, dextrose, glycerol, ethanol or the like and combinations thereof. In addition, if desired, the

composition can contain minor amounts of auxiliary substances such as wetting or emulsifying agents, pH buffering agents and the like which enhance the effectiveness of the active ingredient.

The vaccine composition of the present invention can include pharmaceutically acceptable salts of the components therein. Pharmaceutically acceptable salts include the acid addition salts (formed with the free amino groups of the polypeptide) that are formed with inorganic acids such as, for example, hydrochloric or phosphoric acids, or such organic acids as acetic, tartaric, mandelic and the like. Salts formed with the free carboxyl groups can also be derived from inorganic bases such as, for example, sodium, potassium, ammonium, calcium or ferric hydroxides, and such organic bases as isopropylamine, trimethylamine, 2-ethylamino ethanol, histidine, procaine and the like.

Physiologically tolerable carriers are well known in the art. Exemplary of liquid carriers are sterile aqueous solutions that contain no materials in addition to the active ingredients and water, or contain a buffer such as sodium phosphate at physiological pH value, physiological saline or both, such as phosphate-buffered saline. Still further, aqueous carriers can contain more than one buffer salt, as well as salts such as sodium and potassium chlorides, dextrose, polyethylene glycol and other solutes.

Liquid compositions can also contain liquid phases in addition to and to the exclusion of water. Exemplary of such additional liquid phases are glycerin, vegetable oils such as cottonseed oil, and water-oil emulsions.

An immunocontraceptive composition contains an amount of an sp56 antigenic determinant to as to be effective in inducing an immune response containing anti-sp56 antibodies in the immunized mammal after one or more immunizations. . Routes of immunization can vary, as described further herein, but will affect the formulation of the composition, including the amount of sp56

antigenic determinant required to be effective to induce an immune response. An effective amount, therefore, is any amount which upon administration by a suitable and designated route, induces the production of anti-sp56 antibodies. Repeated immunizations may be utilized, allowing lower concentrations of active ingredient.

Thus, in one embodiment, the present invention provides a method for immunocontraception in a mammal or human comprising administering to the mammal an immunocontraceptive vaccine composition as described herein that contains an sp56 antigenic determinant present on the sperm of the mammal to be treated. The antigenic polypeptide is present in the composition in an amount sufficient to induce an anti-sp56 antibody response following the typical course of administration, which can vary as described herein.

The effectiveness of the vaccine can be determined by observing the presence of anti-sp56 antibodies in serum of the treated mammal, or by measuring the fertility of sperm produces by the treated mammal.

The immunocontraceptive vaccine compositions are conventionally administered intravenously, subcutaneously, intramuscularly or intraperitoneally, but in many preferred embodiments are administered orally in bait or in feed.

The compositions are administered in a manner compatible with the dosage formulation, and in a immunotherapeutically effective amount. The quantity to be administered depends on the subject to be treated, capacity of the subject's system to utilize the active ingredient, and degree of immunoresponsiveness desired. Precise amounts of active ingredient required to be administered can depend on the judgement of the practitioner and are particular to each individual. However, suitable dosage ranges for systemic application are disclosed herein and depend on the route of administration. Suitable regimes for initial administration and booster dosages are also variable, but are typified by an initial administration followed by repeated doses

at one or more hour intervals by a subsequent injection or other administration.

Preferably, an antigenic polypeptide of this invention is present in a composition in an isolated form, i.e., comprising at least about 0.1 percent by weight of the total composition, preferably at least 1%, and more preferably at least about 10%. Particularly preferred for use in formulating compositions is a substantially pure preparation of antigenic polypeptide, i.e., that is initally at least 90% by weight, and more preferably at least 99% by weight prior to admixture in a composition of this invention. Biochemical methods useful for the enrichment and preparation of an isolated polypeptide based on the chemical properties of a polypeptide are well known, and can be routinely used for the production of proteins which are enriched by greater than 99% by weight.

As a prime use of the present invention is the control of pest feral mammal populations, such as rodents and the like, the use of bait or feed is considered particularly useful as a formulation for a vaccine composition of the present invention. As such, a preferred route of administration is the oral route by food ingestion. The composition can include a plant material or other carrier, and therefore, a preferred carrier is a non-toxic, edible carrier. The formulation of baits and feeds into pellets, capsules, tablets, food bars, and the like is generally well known in the agricultural arts, and is not deemed to be limiting to the present invention.

A particularly preferred formulation involves the use of a recombinant DNA expression vector for production of an antigenic polypeptide in a plant, thereby allowing the harvesting of the plant into a feed or bait, thereby providing the vaccine formulation. Exemplary methods include the preparation of a transgenic plant that contains an expressible gene which encodes an antigenic polypeptide of this invention. Alternatively, plants can be infected with viral vectors engineered to contain and express the antigenic polypeptide. A preferred embodiment

described herein uses TMV for infection of plants and display of the antigenic polypeptide in the plant material ultimately used to prepare bait or feed.

G. Anti-sp56 Antibodies

An antibody of the present invention, i.e., an anti-sp56 antibody, in one embodiment is characterized as comprising antibody molecules that immunoreact with an sp56 antigenic determinant of this invention.

Thus, the invention describes an anti-sp56 antibody that immunoreacts with any of the antigenic polypeptides of this invention, and preferably also immunoreacts with native sp56 protein. Preferably, an antibody is substantially free from immunoreaction with a sperm protein other than sp56. Assays for immunoreaction useful for assessing immunoreactivity are described herein.

In one embodiment, antibody molecules are described that immunoreact with a preselected domain of sp56 and not other domains, thereby conferring desirable specificities to the antibody. For example, as shown herein, domains of sp56 have species specific sequences (sss) which are unique to the species of sp56 from which they were derived. Antibodies immunoreactive with a sss domain will not immunoreact with sp56 from a different species, and therefore are useful for typing the sp56 species. In contrast, domains of sp56 are shown herein to be conserved among species, and therefore antibodies immunoreactive with these conserved domains will immunoreact with other sp56 species, and have the advantages of broad species reactivities.

An antibody of the present invention is typically produced by immunizing a mammal with an inoculum containing an antigenic polypeptide of this invention and thereby induce in the mammal antibody molecules having immunospecificity for the antigenic determinant on the immunizing polypeptide and for sp56. The antibody molecules are then collected from the mammal and

isolated to the extent desired by well known techniques such as, for example, by using DEAE Sephadex to obtain the IgG fraction. Exemplary antibody preparation methods using antigenic polypeptides in the immunogen are described extensively in the antibody arts.

The term "antibody" in its various grammatical forms is used herein as a collective noun that refers to a population of immunoglobulin molecules and/or immunologically active portions of immunoglobulin molecules, i.e., molecules that contain an antibody combining site or paratope.

An "antibody combining site" is that structural portion of an antibody molecule comprised of heavy and light chain variable and hypervariable regions that specifically binds antigen.

The phrase "antibody molecule" in its various grammatical forms as used herein contemplates both an intact immunoglobulin molecule and an immunologically active portion of an immunoglobulin molecule.

Exemplary antibody molecules for use in the diagnostic methods and systems of the present invention are intact immunoglobulin molecules, substantially intact immunoglobulin molecules and those portions of an immunoglobulin molecule that contain the paratope, including those portions known in the art as Fab, Fab', $F(ab')_2$ and F(v).

Fab and $F(ab')_2$ portions of antibodies are prepared by the proteolytic reaction of papain and pepsin, respectively, on substantially intact antibodies by methods that are well known. See for example, U.S. Patent No. 4,342,566 to Theofilopolous and Dixon. Fab' antibody portions are also well known and are produced from $F(ab')_2$ portions followed by reduction of the disulfide bonds linking the two heavy reduction of the disulfide bonds linking the two heavy chain portions as with mercaptoethanol, and followed by alkylation of the resulting protein mercaptan with a reagent such as iodoacetamide. An antibody containing intact antibody molecules are preferred, and are utilized as illustrative herein.

The preparation of antibodies against a polypeptide is well known in the art. See Staudt et al., J. Exp. Med., 157:687-704 (1983), or the teachings of Sutcliffe, J.G., as described in United States Patent No. 4,900,811, the teaching of which are hereby incorporated by reference.

Briefly, to produce a peptide antibody composition of this invention, a laboratory mammal is inoculated with an immunologically effective amount of an antigenic polypeptide, typically as present in a vaccine of the present invention. The anti-sp56 antibody molecules thereby induced are then collected from the mammal and those immunospecific for both the immunizing polypeptide and the native sp56 protein, such as the corresponding recombinant sp56 protein, are isolated to the extent desired by well known techniques such as, for example, by immunoaffinity chromatography using the immunizing polypeptide in the solid phase.

To enhance the specificity of the antibody, the antibodies are preferably purified by immunoaffinity chromatography using solid phase-affixed immunizing polypeptide. The antibody is contacted with the solid phase-affixed immunizing polypeptide for a period of time sufficient for the polypeptide to immunoreact with the antibody molecules to form a solid phase-affixed immunocomplex. The bound antibodies are separated from the complex by standard techniques.

The word "inoculum" in its various grammatical forms is used herein to describe a composition containing an antigenic polypeptide of this invention as an active ingredient used for the preparation of antibodies against an sp56 determinant-containing polypeptide. When a polypeptide is used in an inoculum to induce antibodies it is to be understood that the polypeptide can be used in various embodiments, e.g., alone or linked to a carrier as a conjugate, or as a polypeptide polymer. However, for ease of expression and in context of a polypeptide inoculum, the various embodiments of the polypeptides of this

invention are collectively referred to herein by the term "polypeptide" and its various grammatical forms.

For a polypeptide that contains fewer than about 35 amino acid residues, it is preferable to use the peptide bound to a carrier for the purpose of inducing the production of antibodies.

One or more additional amino acid residues can be added to the amino- or carboxy-termini of the polypeptide to assist in binding the polypeptide to a carrier. Cysteine residues added at the amino- or carboxy-termini of the polypeptide have been found to be particularly useful for forming conjugates via disulfide bonds. However, other methods well known in the art for preparing conjugates can also be used.

The techniques of polypeptide conjugation or coupling through activated functional groups presently known in the art are particularly applicable. See, for example, Aurameas, et al., Scand. J. Immunol., Vol. 8, Suppl. 7:7-23 (1978) and U.S. Patent No. 4,493,795, No. 3,791,932 and No. 3,839,153. In addition, a site-directed coupling reaction can be carried out so that any loss of activity due to polypeptide orientation after coupling can be minimized. See, for example, Rodwell et al., Biotech., 3:889-894 (1985), and U.S. Patent No. 4,671,958.

Exemplary additional linking procedures include the use of Michael addition reaction products, di-aldehydes such as glutaraldehyde, Klipstein, et al., J. Infect. Dis., 147:318-326 (1983) and the like, or the use of carbodiimide technology as in the use of a water-soluble carbodiimide to form amide links to the carrier. Alternatively, the heterobifunctional cross-linker SPDP (N-succinimidyl-3-(2-pyridyldithio) proprionate)) can be used to conjugate peptides, in which a carboxy-terminal cysteine has been introduced.

Useful carriers for use in the present invention both in the preparation of an antibody of the present invention and in the preparation of a vaccine composition of the present invention are well known in the art, and are generally proteins themselves. Exemplary of such carriers are keyhole limpet hemocyanin (KLH),

edestin, thyroglobulin, albumins such as bovine serum albumin (BSA) or human serum albumin (HSA), red blood cells such as sheep erythrocytes (SRBC), diphtheria toxin, tetanus toxin, cholera toxin subunits A, B or both, as well as polyamino acids such as poly D-lysine:D-glutamic acid, and the like.

The choice of carrier is more dependent upon the ultimate use of the inoculum and is based upon criteria not particularly involved in the present invention. For example, a carrier that does not generate an untoward reaction in the particular animal to be inoculated should be selected.

The present inoculum contains an effective, immunogenic amount of a polypeptide of this invention, typically as a conjugate linked to a carrier. The effective amount of polypeptide per unit dose sufficient to induce an immune response to the immunizing polypeptide depends, among other things, on the species of animal inoculated, the body weight of the animal and the chosen inoculation regimen is well known in the art. Inocula typically contain polypeptide concentrations of about 10 micrograms (μg) to about 500 milligrams (mg) per inoculation (dose), preferably about 50 micrograms to about 50 milligrams per dose.

The term "unit dose" as it pertains to the inocula refers to physically discrete units suitable as unitary dosages for animals, each unit containing a predetermined quantity of active material calculated to produce the desired immunogenic effect in association with the required diluent; i.e., carrier, or vehicle. The specifications for the novel unit dose of an inoculum of this invention are dictated by and are directly dependent on (a) the unique characteristics of the active material and the particular immunologic effect to be achieved, and (b) the limitations inherent in the art of compounding such active material for immunologic use in animals, as disclosed in detail herein, these being features of the present invention.

Inocula are typically prepared from the dried solid polypeptide-conjugate by dispersing the polypeptide-conjugate in

a physiologically tolerable (acceptable) diluent such as water, saline or phosphate-buffered saline to form an aqueous composition.

Inocula can also include an adjuvant as part of the diluent. Adjuvants such as complete Freund's adjuvant (CFA), incomplete Freund's adjuvant (IFA) and alum are materials well known in the art, and are available commercially from several sources.

The antibody so produced can be used, inter alia, in the diagnostic methods and systems of the present invention to detect sp56 present in a sample such as a tissue section or body fluid sample. A preferred anti-sp56 antibody is a monoclonal antibody.

The phrase "monoclonal antibody" in its various grammatical forms refers to a population of antibody molecules that contain only one species of antibody combining site capable of immunoreacting with a particular epitope. A monoclonal antibody thus typically displays a single binding affinity for any epitope with which it immunoreacts. A monoclonal antibody may therefore contain an antibody molecule having a plurality of antibody combining sites, each immunospecific for a different epitope, e.g., a bispecific monoclonal antibody.

A preferred monoclonal antibody of this invention comprises antibody molecules that immunoreact with an antigenic polypeptide of the present invention as described for the anti-sp56 antibodies of this invention. More preferably, the monoclonal antibody also immunoreacts with recombinantly produced or native, isolated sp56 protein.

A monoclonal antibody is typically composed of antibodies produced by clones of a single cell called a hybridoma that secretes (produces) only one kind of antibody molecule. The hybridoma cell is formed by fusing an antibody-producing cell and a myeloma or other self-perpetuating cell line. The preparation of such antibodies was first described by Kohler and Milstein, Nature, 256:495-497 (1975), the description of which is incorporated by reference. The hybridoma supernates so prepared

can be screened for the presence of antibody molecules that immunoreact with an immunizing polypeptide or with whole sp56 protein.

Briefly, to form the hybridoma from which the monoclonal antibody composition is produced, a myeloma or other self-perpetuating cell line is fused with lymphocytes obtained from the spleen of a mammal hyperimmunized with an antigenic polypeptide containing an sp56 antigenic determinant, such as is present in a preferred polypeptide of this invention. The polypeptide-induced hybridoma technology is described by Niman et al., Proc. Natl. Acad. Sci., USA, 80:4949-4953 (1983), the description of which is incorporated herein by reference.

It is preferred that the myeloma cell line used to prepare a hybridoma be from the same species as the lymphocytes. Typically, a mouse of the strain 129 GlX$^+$ is the preferred mammal. Suitable mouse myelomas for use in the present invention include the hypoxanthine-aminopterin-thymidine-sensitive (HAT) cell lines P3X63-Ag8.653, and Sp2/0-Ag14 that are available from the American Type Culture Collection, Rockville, MD, under the designations CRL 1580 and CRL 1581, respectively.

Splenocytes are typically fused with myeloma cells using polyethylene glycol (PEG) 1500. Fused hybrids are selected by their sensitivity to HAT. Hybridomas producing a monoclonal antibody of this invention are identified using the enzyme linked immunosorbent assay (ELISA) described in the Examples.

A monoclonal antibody of the present invention can also be produced by initiating a monoclonal hybridoma culture comprising a nutrient medium containing a hybridoma that produces and secretes antibody molecules of the appropriate polypeptide specificity. The culture is maintained under conditions and for a time period sufficient for the hybridoma to secrete the antibody molecules into the medium. The antibody-containing medium is then collected. The antibody molecules can then be further isolated by well known techniques.

Media useful for the preparation of these compositions are both well known in the art and commercially available and include synthetic culture media, inbred mice and the like. An exemplary synthetic medium is Dulbecco's Minimal Essential Medium (DMEM; Dulbecco et al., Virol. 8:396 (1959)) supplemented with 4.5 gm/l glucose, 20 mM glutamine, and 20% fetal calf serum. An exemplary inbred mouse strain is the Balb/c.

Other methods of producing a monoclonal antibody, a hybridoma cell, or a hybridoma cell culture are also well known. See, for example, the method of isolating monoclonal antibodies from an immunological repertoire as described by Sastry, et al., Proc. Natl. Acad. Sci. USA, 86:5728-5732 (1989); and Huse et al., Science, 246:1275-1281 (1989).

The monoclonal antibodies of this invention can be used in the same manner as disclosed herein for antibodies of the present invention.

For example, the monoclonal antibody can be used in the therapeutic, diagnostic or in vitro methods disclosed herein where immunoreaction with sp56 is desired.

Also contemplated by this invention is the hybridoma cell, and cultures containing a hybridoma cell that produce a monoclonal antibody of this invention.

H. Diagnostic Methods

The present invention contemplates various assay methods for determining the presence, and preferably amount, of sp56, an sp56 antigenic determinant or an anti-sp56 antibody in a body sample such as a tissue sample, including tissue mass or tissue section, or in a biological fluid sample using a polypeptide, polyclonal antibody or monoclonal antibody of this invention as an immunochemical reagent to form an immunoreaction product whose amount relates, either directly or indirectly, to the amount of an sp56 antigenic determinant or antibody thereto in the sample.

Those skilled in the art will understand that there are numerous well known clinical diagnostic chemistry procedures in which an immunochemical reagent of this invention can be used to form an immunoreaction product whose amount relates to the amount of antigen or antibody in a body sample. Thus, while exemplary assay methods are described herein, the invention is not so limited.

Various heterogenous and homogeneous protocols, either competitive or noncompetitive, can be employed in performing an assay method of this invention.

Sperm having sp56 are considered competent for interaction with the zona pellucida of an egg, and presumptively fertile; sperm lacking sp56 are considered incompetent for interaction and presumptively infertile.

Therefore, one embodiment contemplates a method for assaying the amount of sp56 protein in a sperm-containing sample that utilizes an anti-sp56 antibody to immunoreact with sp56 protein in the sample. In this embodiment, the antibody immunoreacts with sp56 to form an sp56-antibody immunoreaction complex, and the complex is detected indicating the presence of sp56 in the sample. Thus this assay format defines a method for determining fertility in a mammal by detecting the presence of sp56 on the sperm of the mammal.

An immunoassay method using an anti-sp56 antibody molecule for assaying the presence, and preferably the amount, of sp56 in a sample typically comprises the steps of:

(a) Forming an immunoreaction admixture by admixing (contacting) a sample with an anti-sp56 antibody of the present invention, preferably a monoclonal antibody. The sample is typically in the form of a sample in a solid phase such that the immunoreaction admixture has both a liquid phase and a solid phase, and the antibody functions as a detection reagent for the presence of sp56 in the sample.

Preferably, the sample is a sperm-containing sample, such as semen or epididymis tissue, that has been prepared for

immunohistological staining as is well known, although other samples may be adsorbed onto a solid phase, including tissue extracts or body fluid. In that case the adsorption onto a solid phase can be conducted as described for well known Western blot or ELISA procedures.

(b) The immunoreaction admixture is maintained under biological assay conditions for a predetermined time period such as about 10 minutes to about 16-20 hours at a temperature of about 4°C to about 45°C that, such time being sufficient for the sp56 present in the sample to immunoreact with (immunologically bind) the antibody and form an sp56-containing immunoreaction product (immunocomplex).

Biological assay conditions are those that maintain the biological activity of the immunochemical reagents of this invention and the sp56 sought to be assayed. Those conditions include a temperature range of about 4°C to about 45°C, a pH value range of about 5 to about 9 and an ionic strength varying from that of distilled water to that equal to about one molar sodium chloride. Methods for optimizing such conditions are well known in the art.

(c) The presence, and preferably amount, of sp56-containing immunoreaction product that formed in step (b) is determined (detected), thereby determining the amount of sp56 present in the sample.

Determining the presence or amount of the immunoreaction product, either directly or indirectly, can be accomplished by assay techniques well known in the art, and typically depend on the type of indicating means used.

Preferably, the determining of step (c) comprises the steps of:

(I) admixing the sp56-containing immunoreaction product with a second antibody to form a second (detecting) immunoreaction admixture, said second antibody molecule having the capacity to immunoreact with the first antibody (primary) in the immunoreaction product.

Antibodies useful as the second antibody include polyclonal or monoclonal antibody preparations raised against the primary antibody.

(ii) maintaining said second immunoreaction admixture for a time period sufficient for said second antibody to complex with the immunoreaction product and form a second immunoreaction product, and

(iii) determining the amount of second antibody present in the second immunoreaction product and thereby the amount of immunoreaction product formed in step (c).

In one embodiment, the second antibody is a labeled antibody (i.e., detecting antibody) such that the label provides an indicating means to detect the presence of the second immunoreaction product formed. The label is measured in the second immunoreaction product, thereby indicating the presence, and preferably amount, of second antibody in the solid phase.

Alternatively, the amount of second antibody can be determined by preparation of an additional reaction admixture having an indicating means that specifically reacts with (binds to) the second antibody, as is well known. Exemplary are third immunoreaction admixtures with a labeled anti-immunoglobulin antibody molecule specific for the second antibody. After third immunoreaction, the formed third immunoreaction product is detected through the presence of the label.

Methods for detecting the amount or presence of sp56 antigenic determinants can also be useful for determining the species of sp56, and thereby of the sperm, present in a sample where the antibody used is specific form sp56 domains that are species specific, such as the sss domains. Thus, the invention also contemplates a method of identifying the species of a sperm donor or sample by using the above assay with anti-sss antibodies.

Another diagnostic embodiment involves a determination of the contraceptive state of a mammal by measuring whether any anti-sp56 antibodies are present in the circulation or semen.

The presence of anti-sp56 antibodies indicates that the mammal is infertile, or that the fertility is compormised.

Therefore, another embodiment contemplates a method for assaying the amount of anti-sp56 antibody in a body sample that utilizes a sp56 antigenic determinant to immunoreact with an anti-sp56 antibody in the sample. In this embodiment, the antibody in the sample immunoreacts with sp56 to form an sp56-antibody immunoreaction complex, and the complex is detected indicating the presence of anti-sp56 antibodies in the sample. Thus this assay format defines a method for determining infertility in a mammal by detecting the presence of immunocontraceptive anti-sp56 antibodies in the mammal. The method is useful to identify whether the present immunocontraceptive method is effective, and also to identify mammals experiencing an autoimmune response that causes sterility.

An immunoassay method using an sp56 antigenic determinant for assaying the presence, and preferably the amount, of anti-sp56 antibodies in a sample typically comprises the steps of:

(a) Forming an immunoreaction admixture by admixing (contacting) a sample with an antigenic polypeptide containing an sp56 antigenic determinant of the present invention. The sample is typically in the form of a fixed sample in a solid phase such that the immunoreaction admixture has both a liquid phase and a solid phase, and the antigenic polypeptide functions as a detection reagent for the presence of anti-sp56 antibody in the sample.

Preferably, the sample is a blood, serum or plasma- sample, such as semen or epididymis tissue, that has been prepared for immunohistological staining as is well known, although other samples may be adsorbed onto a solid phase, including tissue extracts or body fluid. In that case the adsorption onto a solid phase can be conducted as described for well known Western blot procedures.

(b) The immunoreaction admixture is maintained under biological assay conditions for a predetermined time period such as about 10 minutes to about 16-20 hours at a temperature of about 4°C to about 45°C that, such time being sufficient for the sp56 present in the sample to immunoreact with (immunologically bind) the antibody and form an sp56-containing immunoreaction product (immunocomplex).

Biological assay conditions are those that maintain the biological activity of the immunochemical reagents of this invention and the anti-sp56 antibodies sought to be assayed. Those conditions include a temperature range of about 4°C to about 45°C, a pH value range of about 5 to about 9 and an ionic strength varying from that of distilled water to that equal to about one molar sodium chloride. Methods for optimizing such conditions are well known in the art.

(c) The presence, and preferably amount, of sp56-containing immunoreaction product that formed in step (b) is determined (detected), thereby determining the amount of anti-sp56 antibodies present in the sample.

Determining the presence or amount of the immunoreaction product, either directly or indirectly, can be accomplished by assay techniques well known in the art and as described herein, and typically depend on the type of indicating means used.

The word "complex" as used herein refers to the product of a specific binding reaction such as an antibody-antigen or receptor-ligand reaction. Exemplary complexes are immunoreaction products.

As used herein, the terms "label" and "indicating means" in their various grammatical forms refer to single atoms and molecules that are either directly or indirectly involved in the production of a detectable signal to indicate the presence of a complex. Any label or indicating means can be linked to or incorporated in an expressed protein, polypeptide, or antibody molecule that is part of an antibody or monoclonal antibody composition of the present invention, or used separately, and

those atoms or molecules can be used alone or in conjunction with additional reagents. Such labels are themselves well-known in clinical diagnostic chemistry and constitute a part of this invention only insofar as they are utilized with otherwise novel proteins methods and/or systems.

The labeling means can be a fluorescent labeling agent that chemically binds to antibodies or antigens without denaturing them to form a fluorochrome (dye) that is a useful immunofluorescent tracer. Suitable fluorescent labeling agents are fluorochromes such as fluorescein isocyanate (FIC), fluorescein isothiocyante (FITC), 5-dimethylamine-1-naphthalenesulfonyl chloride (DANSC), tetramethylrhodamine isothiocyanate (TRITC), lissamine, rhodamine 8200 sulphonyl chloride (RB 200 SC) and the like. A description of immunofluorescence analysis techniques is found in DeLuca, "Immunofluorescence Analysis", in Antibody As a Tool, Marchalonis, et al., eds., John Wiley & Sons, Ltd., pp. 189-231 (1982), which is incorporated herein by reference.

In preferred embodiments, the indicating group is an enzyme, such as horseradish peroxidase (HRP), glucose oxidase, or the like. In such cases where the principal indicating group is an enzyme such as HRP or glucose oxidase, additional reagents are required to visualize the fact that a receptor-ligand complex (immunoreactant) has formed. Such additional reagents for HRP include hydrogen peroxide and an oxidation dye precursor such as diaminobenzidine. An additional reagent useful with glucose oxidase is 2,2'-amino-di-(3-ethyl-benzthiazoline-G-sulfonic acid) (ABTS).

Radioactive elements are also useful labeling agents and are used illustratively herein. An exemplary radiolabeling agent is a radioactive element that produces gamma ray emissions. Elements which themselves emit gamma rays, such as ^{124}I, ^{125}I, ^{128}I, ^{132}I and ^{51}Cr represent one class of gamma ray emission-producing radioactive element indicating groups. Particularly preferred is ^{125}I. Another group of useful labeling means are those elements

such as ^{11}C, ^{18}F, ^{15}O and ^{13}N which themselves emit positrons. The positrons so emitted produce gamma rays upon encounters with electrons present in the animal's body. Also useful is a beta emitter, such ^{111} indium or ^{3}H.

The linking of labels, i.e., labeling of, polypeptides and proteins is well known in the art. For instance, antibody molecules produced by a hybridoma can be labeled by metabolic incorporation of radioisotope-containing amino acids provided as a component in the culture medium. See, for example, Galfre et al., Meth. Enzymol., 73:3-46 (1981). The techniques of protein conjugation or coupling through activated functional groups are particularly applicable. See, for example, Aurameas, et al., Scand. J. Immunol., Vol. 8 Suppl. 7:7-23 (1978), Rodwell et al., Biotech., 3:889-894 (1984), and U.S. Pat. No. 4,493,795.

The diagnostic method can also utilize for the detection of an immunoreaction product a specific binding agent. A "specific binding agent" is a molecular entity capable of selectively binding a reagent species of the present invention or a complex containing such a species, but is not itself a polypeptide or antibody molecule composition of the present invention. Exemplary specific binding agents are second antibody molecules, complement proteins or fragments thereof, S. aureus protein A, and the like. Preferably the specific binding agent binds the reagent species when that species is present as part of a complex.

In preferred embodiments, the specific binding agent is labeled. However, when the diagnostic method uses a specific binding agent that is not labeled, the agent is typically used as an amplifying means or reagent. In these embodiments, the labeled specific binding agent is capable of specifically binding the amplifying means when the amplifying means is bound to a reagent species-containing complex.

The diagnostic methods of the present invention can be practiced in an "ELISA" format to detect the quantity of sp56 antigenic determinant in a sample. "ELISA" refers to an enzyme-

linked immunosorbent assay that employs an antibody or antigen bound to a solid phase and an enzyme-antigen or enzyme-antibody conjugate to detect and quantify the amount of an antigen present in a sample. A description of the ELISA technique is found in Chapter 22 of the 4th Edition of Basic and Clinical Immunology by D.P. Sites et al., published by Lange Medical Publications of Los Altos, CA in 1982 and in U.S. Patents No. 3,654,090; No. 3,850,752; and No. 4,016,043, which are all incorporated herein by reference.

Thus, in some embodiments, a polypeptide, an antibody or a monoclonal antibody of the present invention can be affixed to a solid matrix to form a solid support used in the subject diagnostic methods.

A reagent is typically affixed to a solid matrix by adsorption from an aqueous medium although other modes of affixation applicable to proteins and polypeptides can be used that are well known to those skilled in the art. Exemplary adsorption methods are described herein.

Useful solid matrices are also well known in the art. Such materials are water insoluble and include the cross-linked dextran available under the trademark SEPHADEX from Pharmacia Fine Chemicals (Piscataway, NJ); agarose; beads of polystyrene beads about 1 micron (μ) to about 5 millimeters (mm) in diameter available from Abbott Laboratories of North Chicago, IL; polyvinyl chloride, polystyrene, cross-linked polyacrylamide, nitrocellulose- or nylon-based webs such as sheets, strips or paddles; or tubes, plates or the wells of a microtiter plate such as those made from polystyrene or polyvinylchloride.

Examples

The following examples are intended to illustrate but are not to be construed as limiting the specification and claims in any way.

1. Preparation of Purified sp56 Protein Immunogen

Murine sp56 protein was purified with two different approaches from male mice. Retired breeder male mice (HSD:ICR strain; Harlan Sprague Dawley Inc. Indianapolis, IN), 8-18 weeks of age, were killed and the cauda epididymes and vas deferens were transferred to prewarmed (37°C) M199M medium (Bleil and Wassarman, J. Cell Biol., 102:1363-1369 (1986)) containing 5 mM EGTA. Tissue was sectioned with microscissors, and sperm were allowed to swim out for 10 minutes. Sperm were transferred to 9 volumes of M199M and allowed to capacitate 30 minutes at 37°C. Capacitated sperm were purified by centrifugation at 25°C through a Percoll gradient at 5,000 rpm for 7 minutes in a Sorvall SS-34 rotor (Du Pont Instruments, Wilmington, DE). The gradient had been prepared by centrifuging a 39-ml volume of 70% Percoll, 25 mM Tris-HCl, pH 7.4, 150 mM NaCl in a Sorvall SS-34 rotor for 1 hour at 19,000 rpm.

Acrosome-intact and acrosome-reacted capacitated sperm were then collected separately from the gradient by Percoll gradients containing 50 mM Tris-HCl and 1 mM EDTA followed by centrifugation at 35,000 x g for 7 minues at 25°C. The separated sperm populations in suspension were removed from the gradient and then diluted into 8 volumes of medium followed by pelleting at 5,000 rpm in a table top centrifuge. The pellets were resuspended in buffer or medium for treatment in subsequent experiments. Sperm purified in this manner are fully motile and fertile. Capacitated sperm and capacitated, Percoll-purified sperm fertilized 8 out of 10 mouse eggs in a parallel experiment. An average of 4 x 10^7 purified sperm were isolated from each male. Water for all experiments was purified on a Nanopure system (Barnstead, Dubuque, IA).

One method of purifying sp56 was by ZP3-affinity chromatography essentially as previously described (Bleil and Wassarman, Proc. Natl. Acad. Sci., USA, 87:5563-5567 (1990)). Percoll-purified, acrosome-intact sperm from 20 males were extracted in 2 ml buffer A (50 mM Tris-HCl, pH 7.4, 150 mM NaCl,

1 mM $CaCl_2$, 1 mM KCl, and 0.2% hydrogenated Triton X-100) containing 2 mg/ml turkey egg white trypsin inhibitor. Under the conditions described here, UV-transparent, hydrogenated Triton X-100 (Calbiochem Corp., San Diego, CA) extracted 100% of sperm sp56, but only 12% of those sperm proteins solubilized by equimolar Triton X-100. Insoluble material was removed by centrifugation at 14,000 x g for 5 minutes. The sample was filtered through a 0.22 μM filter and loaded onto a ZP3 affinity column (Rainin EP column, 10 μM ZP3; Rainin Instruments, Woburn, MA) that had been constructed as previously described (Bleil and Wassarman, Proc. Natl. Acad. Sci., USA, 87:5563-5567 (1990)). The column was washed with 20 volumes of buffer A and bound protein was eluted with a gradient of 0-8 M urea and 0.2% hydrogenated Triton X-100. Fractions (0.1 ml) were collected and analyzed by 2-D SDS-PAGE.

Typical of many extracellular proteins, sp56, which eluted from a ZP3 affinity column under strongly denaturing conditions, contained intramolecular disulfide bonds. This protein, subjected to 2-D SDS-PAGE, migrated as a doublet having coordinates of 40,000 D (40 kDa) under nonreducing conditions and 56 kDa under reducing conditions.

Densitometric analysis of such gels indicated that sp56 represented only 0.2% of sperm proteins extracted with hydrogenated Triton X-100 (~2.3 pg or 25,000 mol sp56/sperm). In contrast, sp56 was undetectable on 2-D SDS-PAGE of Triton-extracted proteins from Percoll-purified, acrosome-reacted sperm. This result agrees with previous observations, which showed that sp56 was absent from acrosome-reacted sperm (Bleil and Wassarman, Proc. Natl. Acad. Sci., USA, 87:5563-5567 (1990)).

To obtain protein sequence as described above, protein samples containing ZP3 affinity-purified sp56, biochemically purified sp56 (see below), or total Triton-extracted protein from Percoll-purified mouse sperm were first subjected to 2-D SDS-PAGE. Gels were equilibrated 15 minutes in transfer buffer (10% methanol, 1 μM DTT, 100 mM CAPS (Na-[3-(cyclohexylamino)-

propane]-sulfonic acid), pH 10, and proteins were transferred to polyvinyldifluoride protein sequencing membrane (Bio-Rad Laboratories) in that buffer at 4°C for 1 hour at 300 mA in an electroblotting apparatus. The membrane was thoroughly washed in H_20 containing 1 μM DTT, and protein was detected by Ponceau stain containing 1 μM DTT. The stained protein located at the position corresponding to sp56 (40 kDA under nonreducing conditions and 56 kDa under reducing conditions) was sliced out of the membrane, thoroughly washed with 50% methanol and 50% H_20 containing 1 μM DTT, and either digested and extracted with trypsin or directly subjected to solid-phase amino acid sequencing. Tryptic fragments of sp56 separated by reverse-phase HPLC were isolated and subjected to solid-phase amino acid sequencing. Trypsin digestion, reverse-phase HPLC, and amino acid sequencing were performed by The Scripps Research Institute Protein Sequencing Core Facility.

The protein (~30 pmol) subjected to amino acid sequencing was found to have a single NH_2- terminal sequence, the first 12 amino acids of which are DCGPPPLLPFAS (SEQ ID NO 2 from position 1 to 12).

An alternative method of purifying sp56 from sperm was based on a biochemical purification. Percoll-purified, acrosome-intact sperm from 40-50 males were extracted with 2 ml 20 mM Tris-HCl, pH 8.5, 0.5% hydrogenated Triton X-100, and the extract was treated with 1 mM PMSF to inactivate proteases. The extract was centrifuged at 14,000 x g for 5 minutes to remove insoluble material. The supernatant was filtered (0.22 μM filter) and loaded onto a 75 x 7.5-mm DEAE-5PW column (Bio-Rad Laboratories, Richmond, CA). Protein was eluted with a 0-0.5 M NaCl gradient in column buffer (20 mM Tris-HCl, pH 8.5, 0.1% hydrogenated Triton X-100) over 20 minutes at a rate of 0.5 ml/minute. Eluate was analyzed at 280 nm. 0.5-ml fractions were collected and subjected to 2-D SDS-PAGE, and the gels were silver stained to identify the position of sp56 (coordinates 40 kDa [unreduced], 56 kDa [reduced]). Fractions containing sp56 were pooled and

concentrated on a filter (Centricon 30; Millipore Corp., Bedford, MA), and 200 µl were loaded onto a 300 x 7.8-mm BIO-SIL SEC 250 column (Bio-Rad Laboratories). The column was equilibrated and run at 0.2 ml/min with 50 mM Tris-HCl, pH 6.8, 0.1 M NaCl, 0.1% hydrogenated Triton X-100. Eluate was analyzed at 280 nm. 0.2-ml fractions were collected and subjected to 2-D SDS-PAGE to identify the position of sp56. Fractions containing sp56 were pooled, concentrated, and used for further studies. Molecular size determinations in the size-exclusion column were based on protein standards of >300, 150, 44, 17, and <10 kD. In cases where denatured proteins were separated (0.1% SDS replacing 0.1% Triton in the column buffer), standards of 200, 60, and 30 kD were used.

Biochemically purified sp56 eluted from size-exclusion columns at an estimated molecular mass of 110 kDa (range = 90-130 kDa, based on size standards). When total Triton-extracted sperm proteins were subjected to size-exclusion chromatography under nondenaturing (0.1% Triton) conditions, fractions containing sp56 (determined by 2-D SDS-PAGE profile) eluted at 110 kDa. When sperm proteins were subjected to size exclusion chromatography under denaturing (0.1% SDS) conditions, fractions containing sp56 (determined by 2-D SDS PAGE profile) eluted at 46 kDa.

Two lines of evidence further indicate that affinity-purified sp56 and the biochemically purified protein are identical. First, the biochemically purified protein's NH_2 terminal amino acid sequence was identical to that of affinity-purified sp56 as listed above. Second, monoclonal antibodies directed against the biochemically purified protein recognized both biochemically purified protein and affinity-purified sp56. In addition, the biochemically purified sp56 comigrated with affinity-purified sp56 on 2-D SDS gels.

2. Isolation of Murine sp56 cDNA

Biochemically purified sp56 prepared in Example 1 above was proteolytically digested with Endo Arg C, trypsin and

chymotrypsin to generate internal fragments for amino acid sequencing by solid-phase automated Edman-degradation peptide sequencing. Quantitation of the starting material and the amino acids released confirmed that only sp56 sequences were obtained.

From the sp56 peptide fragments, degenerate polymerase chain reaction (PCR) primers were designed for use in amplifying a fragment of sp56 cDNA as described by Bookbinder et al., Science, 269:86-88 (1995). Briefly, the primers were designed according to the methods described by Sambrook et al., Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, 1989 and Lee et al., BioTechniques, 14:191 (1993). Nested PCR was performed on a single-stranded cDNA template made from mouse testes poly(A)+ RNA by using oligo(dT)-primed paramagnetic particles with primers N4 (5'CCNTTYGCNWSNCCNACNAAYCA3') (SEQ ID NO 3) and F5W (5'TARAANACRTARTCNCCRTANGC3')(SEQ ID NO 4), followed by another round of PCR of 1/100 of the original amplification product with primers N5 (5'GCNWSNCCNACNAAYCARYTNTA3')(SEQ ID NO 5) and F5W. The degenerate positions are indicated with the IUPAC convention where N is A, G, C or T; Y is C or T; W is A or T; and R is A or G. The N4 and N5 primers comprise overlapping portions of the chymotryptic peptide 1 while primer F5W corresponds to the chymotryptic peptide 2.

The resultant amplified 1 kilobase (kb) product contained a sequence encoding part of the sp56 NH_2-terminus (amino acids 11 to 18 encoded by PCR primer N5), Endo Arg C peptide 1 (amino acids 271 to 294), Endo Arg C peptide 2 (amino acids 291 to 303), tryptic peptide 1 (amino acids 199 to 205), chymotryptic peptide 1 (amino acids 12 to 18), and chymotryptic peptide 2 (amino acids 342 to 349 encoded by PCR primer F5W). Refer to the nucleotide and amino acid sequence of sp56 listed in SEQ ID NO 1 for location of the primers and peptide fragments described herein.

The full-length sp56 cDNA clone was obtained by screening an oligo(dT)-primer mouse testes Lambda-Zap cDNA library (made in Dr. Bleil's laboratory) with the ^{32}P-labeled, random-primed 1 kb

PCR product obtained above or with ^{32}P-labeled primers specific to the 1 kb PCR product having the sequences 5'ACCTCGCATCACGGCTT3' (SEQ ID NO 6) and 5'GAAGGAAACCACTCGCCA3' (SEQ ID NO 7). The sp56 cDNA clone that was isolated included the entire 1 kb PCR product along with the upstream bases encoding the sp56 NH_2-terminal amino acids 1 through 10 confirming that the clone encoded sp56.

The top strand nucleotide sequence of the murine sp56 cDNA clone is listed in GenBank with the Accession number U17108. The sequence is also listed in SEQ ID NO 1. The sp56 cDNA encodes a 547 amino acid open reading frame that is preceded by a presumptive 32-amino acid signal peptide located between the first in-frame methionine and the experimentally determined NH_2-terminus of sp56 beginning with the three amino acids Asp-Cys-Pro. The encoded amino acid residue sequence of murine sp56 cDNA including the presumptive signal peptide, referred to as "pre-pro sp56" protein, is shown in SEQ ID NO 1 under the cDNA and in SEQ ID NO 2 by itself. The holoprotein, called "pro-sp56", lacking the signal peptide, contains an approximate 6,000 dalton C-terminal peptide that is absent from the mature sp56 protein.

It is now clear that cleavage occurs either naturally on motile sperm, during isolation and treatment of epididymal sperm in nondenaturing detergents, and/or upon binding to ZP3. These observations are of significance for the purposes of the present invention, only insofar as the basic, C-terminal tail of pro-sp56 and nucleic acid encoding this sequence can be used as haptens for immunocontraceptive uses and for development of nucleic acid probes for homology studies, respectively.

Whereas sp56 cDNA encodes a 62,000 dalton polypeptide, from the experimentally determined N-terminus to the end of the cDNA open reading frame, the mature sp56 is a 56 kDa protein. The encoded murine sp56 protein lacks an obvious transmembrane-spanning domain confirming that the protein is a peripheral membrane protein. The murine sp56 protein, referred to as sp56, has now been determined to be a member of a superfamily of protein receptors, called the "Sushi superfamily" based on the

presence of a series of 60 amino acid-long homology repeats, each of which contains ten virtually invariant amino acids. sp56 has seven such domains, the positions of which are indicated in Figure 1 indicated by the underlined sequences where each Sushi domain begins and ends with a cysteine (C) amino acid. A graphic representation of sp56 cDNA and encoded amino acid sequence is given in Figure 2.

Consistent with its membership in the superfamily, sp56 also contains a stretch of 44 amino acids, near the C-terminus, the region of which is unique to the protein and is referred to as the sp56 specific domain. As shown in Figure 1, the stretch of 44 amino acids (amino acids 380 to 423 of the mature sp56 as listed in SEQ ID NO 2), intervenes between Sushi domains 6 and 7. This "sp56 species specific domain" has no detectable homology to any other known protein and contains no cysteine residues, eliminating the possibility of secondary structure determined by disulfide bonds. The sp56-specific domain is composed of highly charged amino acids, indicating that it resides on the protein's exterior and is a likely B cell epitope. Computer simulation did not reveal helical or β-sheet secondary structure. Therefore, due to its unique sequence, probable location on the sp56 surface, and lack of secondary structure, this domain of sp56 represents an excellent candidate as a hapten designed to induce an immune response specific to sp56, and, therefore, mouse sperm and more broadly, rodent sperm.

The therapeutic utility of the sp56-derived compositions of this invention is also supported by the limited tissue distribution of sp56. Expression of sp56, in mouse, was determined to be restricted to spermatogenic cells, and neither the protein nor its mRNA are found in mouse heart, liver, kidney, brain, ovary, or muscle. sp56 protein and mRNA were determined to accumulate in round spermatids although the mRNA was apparently degraded following the round spermatid stage. The protein, however, was transported to the plasma membrane overlying heads of elongated spermatids and spermatozoa. In

addition, murine sp56 is found on sperm that bind to the mouse egg ZP3 protein also present on hamsters but is absent from sperm which do not recognize the mouse egg such as guinea pig and human. Thus, the presence or absence of sp56 on the sperm surface accounts for species specificities of sperm-egg recognition, thereby providing the basis for the immunocontraceptive therapies of this invention.

3. Preparation of Recombinant Murine sp56 Protein and Polypeptide Immunogens

To prepare recombinant immunocontraceptive reagents of this invention, a strain of Tobacco Mosaic Virus (TMV) was created for presentation of a fragment of the sp56 polypeptide on its surface. Mice as well as other, closely related species which consume plant material containing the modified virus develop a mucosal and systemic immune response to the recombinant sp56 immunogen.

The preparation of recombinant sp56 and fragments thereof is preferred for use in this invention as a male mouse contains only about 80 ng sp56 making conventional purification as described in Example 1 economically infeasible. However, because the cDNA encoding sp56 was isolated as described in Example 2, the present invention describes the expression of sp56 or selected portions thereof for preparing recombinant sp56-containing polypeptide immunogens.

Specific regions of the cDNA encoding the murine sp56 protein are selected for preparing recombinant molecules. Preferred recombinant proteins include the 579 amino acid pre-prosp56 and the 547 amino acid prosp56. Other preferred sp56 cDNA regions include those nucleotides that encode the external loop regions of the seven Sushi domains. These domains are stretches of approximately 60 amino acids containing ten virtually invariant amino acids. Figure 3 shows a typical Sushi domain, with the 10 virtually invariant amino acids indicated. Disulfide bonds are formed in the invariant pattern C1-C3 and

C2-C4, creating both a big and a small interconnected looped-loop structure within which a stretch of hydrophobic amino acids is located. The "large loop", composed of a stretch of about 25-30 mostly charged amino acids between C1 and C2 diverges considerably within the superfamily and, because of its charged nature, contributes greatly to the protein receptor function of members of the superfamily (Ichinose et al. J. Biol. Chem., 265:13411-13416 (1990)). sp56 has seven such domains, each containing amino acid regions creating large and small loops, referred to as "ll" and "sl" respectively. Polypeptides derived from these regions in the seven Sushi domains are referred to as "Sh" further denoted by the individual domain 1-7 (e.g., Sh1 to Sh7) and whether the polypeptide is derived from the large or small loop (e.g., Sh1ll or Sh1sl). In addition, Sushi domain 6 is characterized by having an additional cysteine residue within the large loop. As such, the large loop is further separated by the cysteine residues into a first loop (Sh6fl) and a middle loope (Sh6ml), the sequences of each of which are listed in Table 1.

Another additional preferred sp56 cDNA region includes the unique sp56 species specific sequence (referred to as "sss") located between Sushi domains 6 and 7 as described in Example 1. The nucleotide positions of this nonconserved unique region begins at position 1313 and ends at position 1444 as shown in SEQ ID NO 1. The corresponding encoded amino acid residues are located at positions 380 to 423 of SEQ ID NO 1 and as shown in Figure 1. Within this sp56 species specific sequence region, a number of preferred polypeptides are contemplated for use as immunocontraceptives of this invention. These polypeptides are referred to as "sss" further denoted by the polypeptide number (e.g., sss1).

cDNA regions corresponding to selected nucleotide positions for encoding all or various portions of sp56 protein are listed in Table 1 in which the corresponding amino acid residue sequence position (AA) is also indicated. The nucleotide position of the

cDNA fragments and the corresponding encoded amino acid residue positions are respectively derived from SEQ ID NOs 1 and 2, wherein the amino acid residue sequence in SEQ ID NO 2 is also shown with the nucleotide sequence in SEQ ID NO 1. The amino acids are defined from SEQ ID NO 2 rather than SEQ ID NO 1 to avoid confusion with the encoding nucleotide sequences listed in SEQ ID NO 1. The positions are indicated after the colon in the SEQ ID NO; for example, cDNA sss1 has the nucleotide sequence in SEQ ID NO 1 from 1313 to 1444 indicated as SEQ ID NO 1:1313-1444. Accordingly, the encoded polypeptide has the amino acid residue sequence SEQ ID NO 2:380-423. The polypeptide counterparts of the designated cDNAs of immunocontraceptives of this invention are indicated in the specification with a "pp" added to the cDNA designation.

Table 1

Designation	cDNA Position	AA position
pre-pro sp56	1:80-1816	2:-32 to 547
prosp56	1:176-1816	2:1-547
mature sp56	1:176-1606	2:1-477
Sh1	1:176-352	2:1-59
Sh1ll	1:182-271	2:3-32
Sh1sl	1:275-313	2:34-46
Sh2	1:365-538	2:64-121
Sh2ll	1:368-445	2:65-90
Sh2sl	1:449-487	2:92-104
Sh3	1:551-733	2:126-186
Sh3ll	1:554-637	2:127-154
Sh3sl	1:641-679	2:156-168
Sh4	1:746-913	2:191-246
Sh4ll	1:749-829	2:192-218
Sh4sl	1:833-871	2:220-232
Sh5	1:926-1114	2:251-313
Sh5ll	1:929-1027	2:252-284
Sh5sl	1:1031-1075	2:286-300
Sh6	1:1124-1312	2:317-379
Sh6ll	1:1130-1225	2:319-350
Sh6sl	1:1229-1270	2:352-365
Sh6fl[1]	1:1130-1189	2:319-338
Sh6ml[2]	1:1193-1125	2:340-350
Sh7	1:1445-1606	2:424-477
Sh7ll	1:1448-1525	2:425-450
Sh7sl	1:1529-1567	2:452-464
sss1	1:1313-1444	2:380-423
sss2	1:1322-1369	2:383-398
sss3	1:1361-1408	2:396-411
sss4	1:1391-1438	2:406-421

Sh6fl[1] and Sh6ml[2] are respectively the first and middle loops separated by cysteine bonds derived from the large loop of Sushi domain 6.

The selection of these regions outlined in Table 1 facilitates the identification of a recombinant sp56 polypeptide immunogen that is an effective immunocontraceptive based on blocking contraception and not based on antibody titer. Furthermore, selection is favored for polypeptides that may contain a B-cell epitope, enabling potential reversal of the contraceptive effect after cessation of immunization. Thus, in certain embodiments, reversibility is a desirable feature of an immunocontraceptive molecule of this invention.

To prepare the entire prosp56 protein listed in Table 1 as an immunogen for inducing contraception in a mammal, the cDNA sequence encoding the pre-prosp56 protein was inserted into an expression vector for expression in mammalian cells in tissue culture. Briefly, a 2043 bp EcoRI/XhoI fragment from clone 7.1 sp56 cDNA as described by Bookbinder et al., Science, 269:86-89 (1995), the fragment containing the sequence encoding the pre-prosp56 cDNA as listed in Table 1, was inserted into the mammalian expression vector pcDNA3 (Invitrogen, San Diego, CA), linearized with the same restriction enzymes to allow for directional insertion of the sp56 cDNA clone thereby forming pcDNA3-sp56 expression vector.

The expression vector construct provided the following aspects: 1) High-level constitutive transcription of the sp56 cDNA in mammalian cells by upstream promoter sequences from the immediate early gene of the human cytomegalovirus; 2) An additional polyadenylation signal, and a transcription termination sequence from the bovine growth hormone gene which enhances RNA stability; 3) Translation initiation from the first encoded ATG of the sp56 signal peptide (i.e., bases corresponding to 81 to 83 of the signal sequence shown in SEQ ID NO 1); and 4)

Neomycin resistance gene for selection of G418 resistant stable cell lines.

Competent recA(-) E. coli (XL1-Blue, Stratagene, La Jolla, CA) were transformed with the pcDNA3-sp56 construct, and the amplified plasmid was purified from the bacteria using Qiagen (Chatsworth, CA) plasmid purification kits. Purified plasmid DNA was introduced into CHO and COS cells by polycationic lipid transfection (Felgner et al., Proc. Natl. Acad. Sci., USA, 84:7413-7417 (1987)). Silver-enhanced immunogold staining (Cheng et al., J. Cell Biol., 125:867-878 (1994)), using Mab 7C5, detected intracellular sp56 in G418-resistant, cloned COS cells which had been transformed with the sp56 insert. Control cells transformed with the plasmid pcDNA3 containing no insert evidenced no sp56 expression. As expected, expressed sp56, lacking the signal peptide, was secreted from these cells and detected on Western Blots of tissue culture media. The expressed, secreted protein migrated at Mr 62 kDa on SDS-PAGE under reducing conditions, indicating that post-translational removal of the C-terminal tail (Bookbinder et al., Science, 269:86-89 (1995)) did not occur in these cells.

Alternative methods to express the entire pro-sp56 polypeptide for the purpose of oral immunization include, but are not limited to, expression from recombinant Baculovirus in insect cells. The preparation of recombinant pre-prosp56 retaining the signal peptide sequence and mature sp56 lacking both the signal peptide sequence and the C-terminal tail as well as the remaining sp56-derived polypeptides listed in Table 1 are similarly accomplished by cloning of the respective cDNA sequences into an appropriate expression vector. The selection of the expression vector and procedures for insertion thereof are well known to one of ordinary skill in the art.

Once secreted, the recombinant proteins, prosp56 or mature sp56 that are released into medium, are immunopurified with anti-sp56 monoclonal antibodies 7C5, 5F12, 12D2, 7H12, or 4B2 (Cheng et al., J. Cell Biol., 125:867-878 (1994)) or with the

recently isolated anti-pro-sp56 monoclonal antibody 10B6. Purified protein are then mixed with feed to orally immunize mammals.

As described below, in a separate approach, the recombinant sp56 polypeptide immunogens (also referred to as haptens) listed in Table 1 are prepared as part of a chimeric polypeptide molecule composed of the tobacco mosaic virus (TMV) coat protein having the selected sp56-specific domain engineered on the TMV C-terminal tail. Based on the replication of TMV, the assembled virions then present multiple copies of this selected sp56 recombinant fragment that, when consumed in plant material, serves as an immunogen.

For identification purposes, recombinant TMV anchored to a sp56 polypeptide are referred to by the particular polypeptide designation after TMV, e.g., TMV-sss1pp where sss1 indicates the cDNA sequence designation and pp indicates the encoded polypeptide.

Figure 4 summarizes the methodology used to prepare recombinant TMV-sp56 polypeptide fusions. The gene for the TMV coat protein (cp), from TMV strain U1 (common) including flanking TMV DNA sequences, was cloned into a Bluescript plasmid vector (Stratagene) using the restriction enzymes NcoI and KpnI. The TMV gene including the sequence for the TMV cp was derived from the plasmid pU3/12-4 as described by Holt et al., Virol., 181:109 (1991), the disclosure of which is hereby incorporated by reference. These restriction sites are flanked in the plasmid by the Bacteriophage T7 and T3 RNA polymerase promoter sequences as shown in Figure 4.

The construction of a selected sp56 cDNA-TMV chimera is described herein. One of ordinary skill in the art will understand that the teachings described for producing one particular TMV-sp56 chimera are applicable and readily adapted for all the sp56 cDNA molecules listed in Table 1. Thus, to prepare a TMV-sp56 anchored polypeptide based on the sss1 cDNA sequence, the sss1 cDNA (SEQ ID NO 1:1313-1444) is inserted into

the TMV coat protein gene to create a chimeric gene that is then inserted into a separate vector containing the entire TMV gene.

For accomplishing this, a stretch of bases at the 3' end of the TMV coat protein gene are first synthesized (cp-a and cp-b), along with bases encoding sss1 (sss1-a and sss1-b), to create two chimeric PCR primers (sss1-a/cp-a and sss1-b/cp-b). The procedure is identical to that described for preparing a comparable ZP3 expressed molecule (Fitchen et al., Vaccine, 13:1051-1057 (1995)), the disclosure of which is hereby incorporated by reference. The sss1-a nucleotide sequence corresponds to the top strand nucleotide sequence in SEQ ID NO 1 from positions 1322 to 1369. The sss1-b nucleotide sequence is the complementary sequence to that of sss1-a.

The nucleotide sequence, written in the 5' to 3' direction, of the sss1-a/cp-a and sss1-b/cp-b primers respectively are 5'ACCAATGTGACGAACAAGACCTACTTATTCGGTCATGAAGAAAATTCTACTGAGCATGCCATG AAAGGTGTGGGTCCTGCAACTTGAGG3' (SEQ ID NO 8) and 5'GGTCTTGTTCGTCACATTGGTTTCAGATATATTTGGGAAAGTGAATAAAGGCTAATCTATTTT GTAAACCAGAGTCTGCTTGAGAGGTCCAAACCAAAC3' (SEQ ID NO 9). The underlined portions of each primer indicate TMV sequences while the nonunderlined portions indicate the complementary sss1 nucleotide sequence.

Primer sss1-b/cp-b and a primer which specifically anneals to the T7 promoter sequence (T7 primer) having the nucleotide sequence 5'TAATACGACTCACTATAGGGAGA3' (SEQ ID NO 10) are used with the plasmid containing the TMV coat protein sequence to amplify a "left arm" of the cp region contiguous with the sss1 cDNA. Primer sss1-a/cp-a and a primer which specifically anneals to the T3 promoter sequence (T3 primer) having the nucleotide sequence 5'AATAACCCTCACTAAAGGGA3' (SEQ ID NO 11) are used with TMV coat protein plasmid to amplify a "right arm" of the cp region contiguous with the sss. Regions cp-a and cp-b encode contiguous, but opposite strands, of the TMV CP region.

To amplify the above-described regions, the polymerase chain reaction (PCR) is performed as described by Fitchen et al.,

Vaccine, 13:1051-1057 (1995) where the cycle parameters used for the PCR were 95°C, 30 seconds (s)/47°C, 60 s/72°C, 60 s for 25 cycles in the first round followed by a 72°C, 600 s incubation in the second round.

The resultant PCR products containing complementary strands of sss1 each with either the left or right TMV arms were mixed, denatured, then reannealed resulting in a reannealed product that was extended by DNA polymerase as schematically shown in Figure 5. The product of this reaction is composed of an upstream sequence encoding the TMV coat protein, followed by the sss1 cDNA sequence encoding the sp56-specific domain from amino acid residues 380-423 (SEQ ID NO 2), followed by a downstream sequence encoding a stop codon and the remainder of the sequence downstream of the TMV coat protein gene (see Figure 5).

The resultant TMV coat protein encoding the gene containing the sss1 cDNA product, designated TMV-sss1pp, is then re-amplified by PCR, using primers T7 and T3, as described above. The amplified material is then restriction-digested with NcoI and KpnI. The restriction digested PCR fragment is then cloned into the corresponding restriction sites of the TMV cDNA fragment in a derivative of pU3/12-4 yielding a modified full-length cDNA clone of TMV.

The resultant chimeric recombinant plasmid thus contains the modified TMV genome in which the sss1 polypeptide-encoding cDNA is inserted. The chimeric plasmid is then linearized for use in in vitro transcription assays, performed as described by Fitchen et al., Vaccine, 13:1051-1057 (1995). Briefly, T7 RNA polymerase is used to catalyze synthesis of full-length TMV genomes including the modified coat protein gene. Typically, 2 μg of linear plasmid provides sufficient material for inoculating 10-20 leaves.

Following the transcription of TMV-sss1 RNA/polypeptide, the infectious product is rubbed onto spinach plants. The transcription product above is diluted 1/10 in 20 mM sodium phosphate at pH 7.2 and 100 ml is applied to the plants. For

large scale inoculation, crude homogente of infected plants from the primary infections are substituted for the transcription product. This form of infection results in virions that present the sp56-specific sss1 domain on their surface, as part of a chimeric coat protein molecule as shown in Figure 5. Literally thousands of haptens are presented per virion by this method allowing for ingestion of large quantities of immunocontraceptive hapten. The virus itself is infectious and, therefore, economical to grow and useful for making a virtually unlimited supply of antigen. Virus infected plants are raised at The Scripps Research Institute according to USDA Guide Lines. TMV virions are rendered non-infectious prior to use by heating and/or autoclaving of the infected plant material.

Construction of TMV-sp56 chimeric genes for expression of the remaining sp56 polypeptides listed in Table 1 is performed as described above with the exception that the chimeric primer pairs are redesigned with the particular sp56 cDNA nucleotide sequence and complement thereof with the cp a and b primer sequences.

Another alternative recombinant approach for producing recombinant sp56 proteins of this invention for production in plants relies upon the use of shuttle vectors, referred to as pRTL2 vectors, followed by the isolation of a sp56-containing fragment for insertion into a plant expression vector, pGA482. References and discussion thereof for the shuttle vector and expression vector systems are described in Section E.

The shuttle vector pRTL2 was engineered to contain, within a HindIII-isolatable restriction fragment, expression control sequences and nucleotide regions of either cholera toxin α or β chains operatively linked to a particular sp56 encoding cDNA fragment for the expression of cholera toxin-sp56 fusion proteins.

In the pRTL vectors described below, the genes encoding the cholera toxin α and β chains have been separately engineered to be directionally ligated into a shuttle vector for operative ligation with a sp56 cDNA sequence resulting in a construct

encoding a cholera toxin-containing fusion protein. Upon the separate subcloning of the cholera toxin α and β fusion protein-encoding cassettes into an appropriate plant expression vector, the vectors are then used to transform separate plants for subsequent crossing to form a F1 plant containing both expression vectors encoding the two fusion proteins.

Two cholera toxin α chains have been designed along with one β chain gene for facilitating the introduction of expressed fusion proteins into cells and stimulating immune responses therein. An unmodified cholera α chain functions as an adjuvant and a carrier protein. Once internalized into a recipient epithelial cell mediated by complexation with the β chain, the adjuvant aspect of the cholera α chain stimulates antigen presentation on the basolateral surface of the cell. The nucleotide sequence of the α chain has been modified as described below resulting in a amino acid change of an arginine to a lysine at position 7, resulting in the α chain functioning only as a carrier protein as the adjuvant portion was effected by the amino acid alteration. The cholera β chain functions as a carrier protein mediating the binding of the β chain to GM1 ganglioside on the apical surface of epithelial cells. The internalization of either expressed α chain is accomplished by complexing with the expressed cholera toxin β chain protein.

Both α and β chain constructs have been designed for expression of an fusion protein containing a sp56 protein of this invention. For production of an immunocontraceptive fusion protein, either the α or β chain or both are used as fusion proteins.

The pRTL2 vectors have the following elements in common, labeled in Figures 6-9, which are useful as shuttled expression elements to mediate the expression of sp56 in a plant expression vector: 35S promoter (35S Pr) derived from 35S transcript gene of cauliflower mosaic virus; tobacco etch virus untranslated leader sequence (TEV L) that serves as translational enhancer; and 35S polyadenylation sequence (pA-S) of the 3' termius of 35S

transcript gene of cauliflower mosaic virus to direct polyadenylation of the transgene transcript. These elements have been engineered into a single cassette that allows isolation of a restriction digested fragment for subsequent subcloning into a plant expression vector.

The plasmid map of the shuttle vector pRTLCTA-sp56 is shown in Figure 6 in which a HindIII segment has been indicated that contains the expression control sequences and illustrates the location of a sp56-encoding DNA fragment (labeled as "sss") operatively linked within the cholera toxin α chain (labeled as CTOX A). To generate a cholera toxin α chain-sp56 fusion construct, a sp56 cDNA sequence is selected from Table 1 and is produced by synthesis or PCR incorporating ClaI restriction sites allowing for ligation into the ClaI site present in the cholera toxin α chain gene. Once the insertion of the sp56-encoding sequence is verified by restriction analysis, the fusion protein-encoding cassette is isolated by HindIII digestion for subsequent ligation into a similarly digested pGA482 plant expression vector. The nucleotide sequence of the cholera toxin α chain fusion protein cassette without a sp56 insert is listed in SEQ ID NO 12 having 1956 bases. The modified cholera toxin α chain encoding a lysine instead of an arginine at position 7 of the cholera toxin gene contains two adenine nucleotides at positions 787 and 788 rather than cytosine and guanine as shown in SEQ ID NO 12. The vector labeled pRTLCTAm-sp56 having the modified α chain is schematically shown in Figure 7, in which all the other vector elements are the same as those described for Figure 6.

The cholera β chain fusion protein construct is shown in Figure 8 also containing the same vector elements as previously described but having the nucleotide sequence encoding cholera toxin β chain. sp56-encoding sequences, containing BglII-specific ends, are inserted in the BglII site within the cholera toxin β chain as indicated by "sss" in the figure. The nucleotide sequence of the cholera toxin β chain HindIII fusion cassette lacking a sp56 insert is listed in SEQ ID NO 13.

In order to express larger sp56 polypeptides and fusion proteins, the vector pCTA02-1sp56, as shown schematically in Figure 9, was also constructed. The nucleotides in the cholera toxin α chain except for the ClaI restriction site was removed thereby allowing for directional ligation of a similarly prepared sp56 insert into a KpnI/ClaI digested vector. As with the other vectors, the fusion protein-expression cassette is isolated by digestion with HindIII for subcloning into a plant expression vector for expression thereof.

Depending on the selected shuttle vectors used for preparing a fusion construct, the HindIII-isolated fragments containing the control elements with the cholera chain-encoding genes/sp56 coding region, or with the one lacking cholera toxin α chain (pCTA02-1sp56), are then separately inserted into a similarly digested pGA482 (An et al., Plant Molecular Biology Manual, A3:1-19 (1988). The recombinant plant vector is then used to transform a plant virus that is then used to infect plants, such as tobacco or alfalfa. A preferred plant virus is Agrobacterium strain LBA4404.

4. Preparation of Synthetic Murine sp56 Polypeptides and Chimeric sp56 Polypeptide Immunogens

In an alternative but equally effective method for preparing immunocontraceptive molecules of this invention, sp56-derived polypeptides are synthesized for use as haptens in conjunction with a selected carrier molecule to effect an immune response when injected into a recipient subject.

As one example of this approach, a 16 amino acid peptide of the sp56 specific sequence, encoded by the sss3 cDNA as listed in Table 1, having the underlined single letter amino acid residue sequence CNISETNVTNKTYLFGH (SEQ ID NO 14), was used as a hapten to orally immunize mice. The peptide was synthesized by The Scripps Research Institute Core Facility and contained an additional cysteine at the N-terminus. The underlined portion of the peptide corresponds to amino acid residue positions 396 to 411 of the sp56 protein listed in SEQ ID NO 2. A peptide having

the amino acid residue sequence CHGFLYTKNTVNTESIN (SEQ ID NO 15), reversed from the sp56 peptide above, was used as a control to assess specificity of the immune response.

This hapten (1.5 mg) was crosslinked to 1 mg Sulfo-SMCC-modified Cholera Toxin B subunit (CT-B, maximum 10 SMCC residues per subunit) via the exposed N-ethyl maleimide. The resultant immunogen was designated CT-B-sss3pp. Cholera Toxin was chosen as the adjuvant, or carrier, because it has been successfully used to present other haptens in order to develop a strong mucosal immune response in the absence of oral tolerance (Black et al., Infect. Immunity, 55:1116-1120 (1987); McGhee et al., Reprod. Fertil. Devel., 6:369-379 (1994)). Alternative reagents, including diptheria toxin, are also useful as a carrier.

CT-B bearing the peptide hapten (CT-B-sss3pp) was purified by size exclusion chromatography and modification confirmed by SDS-PAGE analysis of aliquots. The resultant immunogen was then used to immunize mice as described in Example 5.

The remaining sp56 polypeptides listed in Table 1 are similarly prepared as immunogens for use in inducing immunocontraception upon oral ingestion. One of ordinary skill in the art also understands that the synthetic polypeptides are used with and without conjunction with a carrier molecule and that the carrier molecule is not limited to cholera toxin.

5. Immunization of sp56 Immunogens into Mice

The ability of the sp56 compositions to act as immunocontraceptives was first assessed with purified, denatured, reduced sp56 prepared as described in Example 1. For the test, one male mouse, 25 days of age was immunized intraperitoneally with purified, denatured, reduced sp56 (500 ng in Complete Freund's Adjuvant on day 1 and 500 ng in Incomplete Freund's Adjuvant on day 60). Sham injection of Complete and Incomplete Freund's Adjuvant was performed on another male mouse.

The section of a testicle from the male which had been immunized with sp56 was analyzed immunohistochemically using goat anti-mouse Ig conjugated Horseradish Peroxidase, developed with Diaminobenzidine, and counterstained with Haematoxylin/Eosin. When compared with a section of a testicle from the sham immunized male, immunohistochemically stained in the same way, the sp56-immunized male had been subject to removal of elongating spermatids and spermatozoa. Clearing of late stages of spermatogenesis that was seen was most likely due to immune attack, since atretic cells were positively stained for mouse Ig. The immunized male copulated with and plugged four female mice, but no offspring were observed. The sham immunized male produced five litters. This result demonstrates that immune response to sp56, in mice, results in male infertility.

To assess other sp56-derived polypeptides in acting as immunocontraceptives, hapten-modified CT-B prepared in Example 4 (CT-B-sss3pp, experimental, mice #'s 5-8) or CT-B alone (CT-B, control, mice #'s 1-4) was administered to 28 day old virgin male mice (HSD:ICR strain, Harlan Sprague-Dawley) by weekly intubation, using approximately 20 μg protein in 100 μl saline per week per mouse. After three weekly intubations, serum was obtained by intraorbital bleed and mucosal extract was obtained from homogenized fecal pellets. Diluted serum (50 μl, 1/10, in Tris-buffered saline) and mucosal extract (50 μl from one fecal pellet homogenized in 0.5 ml Tris-buffered saline) were incubated in 96-well plates containing purified, lyophilized sp56 (5 ng) and analyzed for antigen recognition in modified ELISA, performed as previously described (Cheng et al., J. Biol. Chem., 125:867-878 (1994)). Silver-enhanced colloidal gold-conjugates of goat anti-mouse IgG (Sigma, St. Louis, MO) and goat anti-mouse IgA (Sigma) were used to detect bound serum and mucosal Ig's, respectively.

Mice immunized with the sp56-derived hapten (CT-B-sss3pp, experimental) contained antibodies which recognized purified sp56 as shown in Table 2. The results were identical for both serum

IgG and mucosal IgA. As such, only the results, presented as positive (+) or negative (-) are shown for each immunogen.

Table 2

Mouse #	Immunogen	Results
1	CT-B	-
2	CT-B	-
3	CT-B	-
4	CT-B	-
5	CT-B-sss3pp	+
6	CT-B-sss3pp	+
7	CT-B-sss3pp	+
8	CT-B-sss3pp	+

As shown in Table 2, all four CT-B-sss3pp immunized mice contained both IgA and IgG which bound to sp56, whereas none of the four CT-B immunized mice contained antibodies which detectably recognized the protein.

To characterize the antiserum of immunized mice, both immunohistochemical and Western blot analyses were respectively performed on mouse sperm and solubilized mouse sperm proteins.

Serum from immunized mice was obtained by intraorbital bleed after three weekly oral immunizations. For the immunohistochemical analyses, diluted serum (1/10 in Tris-buffered saline) was tested for IgG recognition of fixed, acrosome-intact mouse sperm (from nonimmunized males) on glass slides by silver-enhanced immunogold staining, using colloidal gold conjugated goat anti-mouse IgG (Sigma), as previously described (Cheng et al., J. Biol. Chem., 125:867-878 (1994); Bookbinder et al., Science, 269:86-89 (1995)).

IgG from mouse #5, immunized with CT-B-sss3pp, bound exclusively to a region of plasma membrane overlying the nonimmunized sperm's acrosome, where sp56 is located. Serum IgG from mice #'s 6-8 gave the same pattern of immunohistochemical

staining. Binding of IgG from mouse #1, immunized with CT-B alone, was not detected. Serum IgG from mice #'s 2-4 gave the same pattern of immunohistochemical staining.

Serum Ig's from mice #'s 1 and 5 (immunized with control CT-B and CT-B-sss3pp, respectively) were analyzed for antigen specificity by Western blots. For this assay, total, Triton-soluble proteins from sperm of nonimmunized mice were subjected to SDS-PAGE, transferred to Immobilon (Millipore) membrane, and strips (approximately containing 505 grams of sperm protein per strip) were probed with immune serum as previously described (Cheng et al., J. Biol. Chem., 125:867-878 (1994)). The antibody-probed membrane was developed by silver-enhanced immunogold staining as previously described. Primary antibodies used to probe the blots were serum (1/100 dilution); from mouse #1 or from mouse #5 (1/10 dilution) or anti-sp56 monoclonal antibody 7H12 (50 μg/ml; positive control).

Western Blot analysis of total, detergent-soluble sperm proteins indicated that sp56 was the only protein recognized by anti-sss antibodies. Serum IgG from CT-B-sss3pp-immunized mice recognized a single Mr 40 kDa band on nonreduced gels as shown in Figure 10 where the positive control monoclonal antibody 7H12 also immunoreacted with the 40 kDa protein. The "o" indicates the origin of the separating gel. These results demonstrate that the sp56 sss3pp hapten chosen was successfully used to orally immunize mice, and that the immune response was directed against sp56 in situ.

Sperm from immunized mice were then analyzed by immunohistochemistry to assess the direct effects of the sp56 polypeptide immunocontraceptive compositions of this invention. For this assay, cauda epididymal sperm were isolated from immunized mice and incubated for 30 minutes in a medium which supports capacitation, as previously described (Bleil and Wassarman, Proc. Natl. Acad. Sci., USA, 85:6778-6783 (1988)). Incubated sperm were fixed and prepared for immunohistochemistry as previously described above. Dried, plated sperm were

rehydrated and incubated for 4 hours at 25°C in TNBG (50 mM Tris-HCl, pH 7.5, 150 mM NaCl, 4 mg/ml Bovine Serum Albumin (BSA) [Sigma], and 5% normal goat serum [Sigma]), followed by a 4 hour incubation in TNBG containing 50 μg/ml colloidal gold-conjugated goat anti-mouse IgA (Sigma). Sperm were washed, silver-enhanced, and counterstained as previously described.

Gross observation did not indicate a difference between levels (approximately 70%) or types of motility of sperm from control or experimental mice. Acrosome-intact sperm from CT-B-sss3pp-immunized mice evidenced immunogold staining restricted solely to the surface of plasma membrane overlying the acrosome, a "crescent" defining the location of sp56. No immunogold staining was detected on sperm from control CT-B immunized mice.

Sperm from immunized mice are then characterized in a sperm-egg binding assay. A portion of the sperm incubated for capacitation as described above were challenged with cumulus-free mouse eggs in a sperm-egg binding assay, as previously described (Bleil and Wassarman, Proc. Natl. Acad. Sci., USA, 85:6778-6783 (1988)). Numbers of sperm bound per egg (10 eggs and 2 two-cell embryos per assay) are measured as previously described.

Other mice, both male and female, are immunized with comparably prepared sp56 immunocontraceptive compositions prepared as described in Examples 3 and 4 respectively for recombinant and synthetic molecules. The response to immunization is assessed as described herein for each of the assayed sp56 protein or polypeptide reagents that are prepared either in recombinant form or by synthesis with and without carrier molecules.

Male mice that have been successfully immunized with an sp56 immunocontraceptive protein or polypeptide of this invention are then mated to evaluate reproductive capabilities. All males and females from groups containing at least one member evidencing an immune response by ELISA are singly mated with nonimmunized partners on day 60 (60 days post-immunization) to test for

infertility. Couples are caged together for a maximum of 20 days and females are checked for copulation plugs each morning. Plugged, nonimmunized females are checked for pregnancy, both by gross observation of belly distention and by parturition, during the 20 day period and for an additional 21 days following that period. Success of contraception of immunized males are measured as percent of nonimmunized females (plugged by those immunized males) which fail to give birth. Plugged, immunized females are separated and checked for pregnancy in the same way. Success of contraception of immunized females is measured as percent of plugged, immunized females which fail to give birth. The minimal requirement for success rate for any group is 50% contraception for males and 50% for females. Even with these low numbers, the reagent efficiency is sufficient to depress a population of rodents.

On day 101, 10-20 successfully immunized, contraceptive males (no more than 5 from any group) are then analyzed to determine the site(s) of immunocontraception. These males are mated with nonimmunized females. Plugged females are sacrificed 7 hours following copulation and their oviducts and uterine horns dissected and flushed with saline. Total cellular material is isolated by centrifugation. Soluble material in flushes is analyzed for anti-sp56 Ig by ELISA as described above. Ig typing is performed by ELISA. Absence of sperm in flushes indicates immune attack and clearing of sperm upstream of the vas deferens. Sperm found in flushes is observed for impaired motility, association with T cells or macrophages, and antibody coating using Goat anti-mouse Ig-Horseradish Peroxidase conjugate. In all cases, Fc receptor binding is pre-blocked with normal goat serum, as previously described (Cheng et al., J. Biol. Chem., 125:867-878 (1994)). Presence of sperm and association of Ig and/or immune cells indicates mucosal immune attack on differentiated sperm in the epididymus or in semen.

Half of the successfully immunized, infertile males are subsequently sacrificed and reproductive organs analyzed for

immunization-dependent pathology. Testes and epididymides are fixed, embedded, sectioned, and stained as described above. Testicular and epididymal sections of these infertile, sp56-immunized males are analyzed by immunohistochemistry to define the method of immune attack, as described above. This procedure is used to determine whether the immune response involves invasion by cells of the immune system (e.g., macrophages, killer T cells) or simply by attack by circulating or secreted antibodies. Carcasses of these 5-10 sp56-immunized males are further analyzed for anatomical pathology in other organs by the Scripps Animal Facility Veterinarian.

If sp56-immunized females evidence contraception (i.e., successful copulation, as determined by plugs, but no offspring), they are then tested for immune attack upon sperm in oviducts and uterine horns. As many as ten contraceptive, TMV-sp56 immunized females are mated with nonimmunized males and sperm within their oviducts and uterine horns analyzed for the basis of immune attack, as described above.

An additional assay is performed to determine whether TMV-sp56 protein or polypeptide immunization is reversible. Reversibility of sp56-dependent infertility may be desirable, if threatened or endangered rodents occupy the same niche as nuisance rodent populations. For this aspect, the fertility of the remaining, successfully immunized males and females is followed. After day 81, these animals are be fed food pellets containing wild-type TMV or other delivery vehicle. For a period of 120 days, their fertility is tested as described above. Success of reversibility of sp56-dependent infertility is measured as percent of non-immunized females giving birth after copulation with sp56-immunized males and percent of sp56-immunized females giving birth after copulation with nonimmunized males.

6. Isolation of Mammalian cDNA Homologs of Murine sp56

Mammalian cDNA homologs of the murine cDNA sp56 obtained as described in Example 2 are obtained by the screening protocol described herein. While the protocol is designed for screening any mammalian tissue, the procedure is described for isolating only the human homolog. One of ordinary skill in the art understands that the methods described for isolating the human sp56 homolog are readily extrapolated to use in identifying other mammalian homologs. Furthermore, one skilled is the art also understands that the identification of any mammalian homolog is not limited to the methods described herein.

In preliminary screening, low stringency Northern Blot analysis of human testicular RNA, probed with the full-length cDNA encoding mouse sperm sp56, detected faint RNA bands in the 1-2 kb range. Because hybridizing RNA was only detected at low stringency, a homolog encoded by that RNA is probably distantly related to sp56. This might be expected, since the human oocyte ZP recognition domain of the human sperm protein may recognize ZP polypeptide, instead of carbohydrate, as is the case in mouse (Florman et al., Devel. Biol., 106:2430151 (1984); Florman and Wassarman, Cell, 41:313-324 (1985); Bleil and Wassarman, Proc. Natl. Acad. Sci., USA, 85:6778-6783 (1988)).

Additional Northern Blot analysis of human testicular RNA is performed in order to identify probes (derived from sp56 cDNA) for identifying human testicular cDNA's encoding sp56 homologs. All Northern Blot analyses are tested for successful RNA isolation, transfer, and presentation by post-screening with a probe for G3PDH mRNA. Two types of nucleic acid probes are used to detect human testicular RNA encoding sp56 homologs on Northern Blots. Those probes providing the highest peak-to-noise ratio and which clearly bind to the smallest number of RNA bands on blots are chosen for cDNA screening.

Those probes are described as follows:

Probes Type I: Double-stranded ^{32}P-labeled random-primed DNA sp56 cDNA probes are prepared from the full-length sp56 cDNA clone 7.1

having the nucleotide sequence in SEQ ID NO 1, or from subclones of the cDNA, and used to probe Northern blots as previously described (Bookbinder et al., Science, 269:86-89 (1995). The full-length probe has the advantage of having been successfully used to identify sp56 RNA in mouse testicular RNA Northern Blots and of detecting, at relatively low stringency, human testicular RNA's. Additional Northern Blot studies with the full-length probe are performed, where stringency of post-hybridization wash conditions is varied in order to obtain better peak-to-noise ratios. For example, post-hybridization washes are repeated using 4X SSC, 0.5% SDS at 25°C, and concentrations of SSC subsequently varied (4X to 6X), concentrations of SDS varied (0.5% to 0.1%), and blocking agents, such as Denhardt's Reagent added in some experiments. Alternatively, other probes, representing sub-clones of the sp56 cDNA, and corresponding to the 5' half of sp56 cDNA clone 7.1 with bases 1 through 1300 in SEQ ID NO 1 or to the 3' half of sp56 cDNA clone 7.1 with bases 1301 through 2050 (SEQ ID NO 1) are used, as described above, to screen Northern Blots.

Probes Type II: Single-stranded, synthetic biotinylated deoxyribonucleotide 60-mer antisense probes derived from sp56 cDNA, bearing biotin at the 3' end, are synthesized by The Scripps Institute Core Facility. These antisense probes have the following described complementary sequences written in the 5' to 3' direction to the region indicated in the sp56 cDNA sequence listed as Genbank accession number U17108 and in SEQ ID NO 1:

5'TCTGAATAAGACGTAGGAAGACTCTAGTGTTGCCTTGGGTTTTTCGCCCCTCCTCGTGCC3' (SEQ ID NO 16) complementary to nucleotides 1-60 (SEQ ID NO 1) in the 5' untranslated region;

5'GTTGTTGACTCATACAACTGGTTGGTTGGAGAAGCAAACGGTAAAAGGGGTGGAGGTCCA3' (SEQ ID NO 17) complementary to nucleotides 181 to 240 (SEQ ID NO 1) in Sushi domain 1;

5'CCCTTCTTGCATCGGATCTCAAGAGCGTCTCTGTACGCATAGAACTGTCGAAATCCAGGG3' (SEQ ID NO 18) complementary to nucleotides 781-840 (SEQ ID NO 1) in Sushi domain 4; and

5'CCTTTCATGGCATGCTCGGTTGAGTTTTCTTCATGACCAAATAGATAGGTTTTGTTTGTC3' (SEQ ID NO 19) complementary to nucleotides 1381-1440 (SEQ ID NO 1) in the sequence-specific domain (sss).

The four probes were chosen on the basis of anticipated sp56 homology to possible human sperm analogues of sp56. The 5' untranslated region was chosen because of its probable regulatory function in expression. The regions of Sushi domains 1 and 4 are largely conserved, under the indication that the N-terminal half of sp56 represents a conserved sperm anchoring domain. The sp56-specific sequence (sss), as part of the carbohydrate recognition domain of sp56, is not likely conserved, and, in Northern Blot analysis, functions as a negative control. However, any one (or all) of the probes which clearly and satisfactorily detect human testicular RNA on Northern Blots are subsequently used for cDNA screening. Single-stranded, biotinylated polynucleotide probes are hybridized to Northern blots containing UV-crosslinked RNA from human testes and from mouse testes (positive control) under conditions suggested by the manufacturer (Millipore, BIORAD). Post-hybridization washes are performed under varying levels of stringency, in order to obtain the highest peak-to-noise ratios. Blots are subsequently probed for detection of hybrids, using streptavidin-alkaline phosphatase conjugates and developed using BCIP-BPT reagent.

sp56 cDNA sequences, probes of which have been shown to successfully detect distinct RNA bands on Northern blots bearing human testicular RNA, are then used to screen a human testicular cDNA library. Dr. Bleil's laboratory has developed an oligo dT-primed human testicular cDNA library, created as previously described (Bookbinder et al., Science, 269:86-89 (1995)), and containing approximately 10^8 independent clones in the Lambda-Zap II vector. Comparable libraries are also commercially available. Probes selected as described above are used to screen the library essentially as previously described for screening of a mouse testis cDNA library. Plaques detected in such screens are cloned and cDNA's sequenced as previously described.

Sequenced cDNA's are identified as candidates for the cDNA which encodes the human analogue of sp56 based on the following criteria:

1. The probes which identify cDNA clones must also detect distinct RNA bands on Northern Blots of both mouse (positive control) and human testicular RNA. As described above, probes encoding the sss of sp56 are not anticipated to identify human testicular RNA's or their respective cDNA's.

2. The translated sequence of selected cDNA's must have considerable homology to (at least half of) the sp56 amino acid sequence. Since sp56 is a member of a superfamily of protein receptors, each of which contains short homology repeats, called Sushi Domains, other members of the superfamily, which are not sp56 homologs, may be inadvertently identified. For example, the alpha subunit of complement 4B-binding protein (C4BP) is most closely related to sp56, having overall amino acid identity of 47%. However, the N-terminal half of sp56, from Sushi domains 1 through 5, has 54% amino acid identity to human C4BP, whereas the C-terminal half, containing Sushi domain 6, the SSS, Sushi domain 7, and the basic C-terminal tail, has only 36% amino acid identity to human C4BP. Thus, in order to be a candidate as the human sperm homolog of sp56, the translated sequence of selected cDNA's must have either an overall amino acid identity to sp56 of greater than 47%, or the N-terminal half must have greater than 54% identity, or the C-terminal half must have greater than 36% identity. These requirements have been chosen in order to minimize the chance of identifying proteins which are not functional homologs of mouse sperm sp56.

Selected, sequenced cDNA's, meeting the above criteria, are identified as encoding candidates for the human sperm homolog of sp56. A human sp56 homolog must be restricted to sperm and to late stages of spermatogenesis and be absent from other human tissues. In order to test whether expression of candidate proteins is restricted in human, Northern Blot analysis of RNA extracted from human tissues (heart, liver, epididymis, testes,

muscle, lung, brain, kidney, and ovary are performed as previously described using ^{32}P hybridization probes derived from selected human testicular cDNA's. If, like murine sp56, Northern Blots reveal that RNA encoding candidate sp56 homologs is expressed exclusively in testis, the encoded protein represents an ideal immunocontraceptive target in human, regardless of its function in vivo. One other characteristic required of such a candidate, if it is to be an effective target, is that the protein is expressed on the surface of human sperm.

The identification and sequence analysis of a human or other mammalian homolog then allows for the preparation of sp56 protein and polypeptide immunogens prepared essentially as described in Examples 3 and 4 and delivered as described in Example 5.

The foregoing specification, including the specific embodiments and examples, is intended to be illustrative of the present invention and is not to be taken as limiting. Numerous other variations and modifications can be effected without departing from the true spirit and scope of the invention.

SEQUENCE LISTING

(1) GENERAL INFORMATION:

(i) APPLICANT:

(A) THE SCRIPPS RESEARCH INSTITUE

(B) STREET: 10550 North Torrey Pines Road

(C) CITY: La Jolla

(D) STATE: California

(E) COUNTRY: United States

(F) ZIP: 92037

(G) TELEPHONE: (619) 554-2937

(H) TELEFAX: (619) 554-6312

(ii) TITLE OF INVENTION: IMMUNOCONTRACEPTION COMPOSITIONS CONTAINING SPERM ANTIGEN, AND METHODS OF USE

(iii) NUMBER OF SEQUENCES: 19

(iv) COMPUTER READABLE FORM:

(A) MEDIUM TYPE: Floppy disk

(B) COMPUTER: IBM PC compatible

(C) OPERATING SYSTEM: PC-DOS/MS-DOS

(D) SOFTWARE: PatentIn Release #1.0, Version #1.25

(v) CURRENT APPLICATION DATA:

(A) APPLICATION NUMBER: PCT/US97/

(B) FILING DATE: 03-JUN-1997

(vi) PRIOR APPLICATION DATA:

(A) APPLICATION NUMBER: US 08/674,895

(B) FILING DATE: 03-JUN-1996

(2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 2026 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(ix) FEATURE:

(A) NAME/KEY: sig_peptide

(B) LOCATION: 81..175

(ix) FEATURE:

(A) NAME/KEY: mat_peptide

(B) LOCATION: 176..1816

(ix) FEATURE:

(A) NAME/KEY: CDS

(B) LOCATION: 80..1819

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

```
GGCACGAGGA GGGGCGAAAA ACCCAAGGCA ACACTAGAGT CTTCCTACGT CTTATTCAGA   60
TACCTACAGA AAAGGGAGA ATG ATA ACC TGG TCC TTC ATT GAT CTG TGG AGA   112
                     Met Ile Thr Trp Ser Phe Ile Asp Leu Trp Arg
                     -32     -30                 -25
ACC TCT CAT TCA ACT CTG TTC CAA ATG ACC TTG GCC ACT GTT CTG ATG    160
Thr Ser His Ser Thr Leu Phe Gln Met Thr Leu Ala Thr Val Leu Met
    -20                 -15                 -10
GCT CCT GTT CTT GGT GAT TGT GGA CCT CCA CCC CTT TTA CCG TTT GCT    208
Ala Pro Val Leu Gly Asp Cys Gly Pro Pro Pro Leu Leu Pro Phe Ala
 -5                  1               5                  10
TCT CCA ACC AAC CAG TTG TAT GAG TCA ACA ACC TTC CCA TCT GGA ACT    256
Ser Pro Thr Asn Gln Leu Tyr Glu Ser Thr Thr Phe Pro Ser Gly Thr
             15                  20                  25
GTC CTG AAA TAT ACC TGC CAT CAC GGC TTC AAG AGA GTC AAT TCA AGC    304
Val Leu Lys Tyr Thr Cys His His Gly Phe Lys Arg Val Asn Ser Ser
         30                  35                  40
CAT CTT TCT TGT GAT GAG AAT GGT TCA TGG GTC TAT AGT ACC TTT TGT    352
His Leu Ser Cys Asp Glu Asn Gly Ser Trp Val Tyr Ser Thr Phe Cys
     45                  50                  55
GCC AGG AAA CGA TGC AAG AAC CCA GGC GAG TTG GTC AAT GGG AAA GTA    400
Ala Arg Lys Arg Cys Lys Asn Pro Gly Glu Leu Val Asn Gly Lys Val
 60                  65                  70                  75
GAA ATT CCA TCT GAC CTT TTG GTA GGC TCA ATC ATA GAG TTC AGC TGC    448
Glu Ile Pro Ser Asp Leu Leu Val Gly Ser Ile Ile Glu Phe Ser Cys
                 80                  85                  90
TCA AAG GGC TAT CTT CTG ATT GGC TCA GCA ACT AGT CGG TGT GAG GTC    496
Ser Lys Gly Tyr Leu Leu Ile Gly Ser Ala Thr Ser Arg Cys Glu Val
             95                  100                 105
```

```
CAA GGT AAA GGA GTT GAC TGG AGT GAT TCT CTC CCA GAA TGT GTA ATT     544
Gln Gly Lys Gly Val Asp Trp Ser Asp Ser Leu Pro Glu Cys Val Ile
        110                 115                 120
GCC ACG TGT GAG CCC CCT CCG CCC ATC AGC AAT GGG AAG CAC AGT GGG     592
Ala Thr Cys Glu Pro Pro Pro Pro Ile Ser Asn Gly Lys His Ser Gly
    125                 130                 135
AGA GAT GAT GAC CTG TAC ACG TTT GGC TCT GTA GTC ATC TAC AAT TGT     640
Arg Asp Asp Asp Leu Tyr Thr Phe Gly Ser Val Val Ile Tyr Asn Cys
140                 145                 150                 155
GAT CCC ACC TTC ACA CTC CTT GGC AAT GCC TCC ATT GTC TGC ACT GTG     688
Asp Pro Thr Phe Thr Leu Leu Gly Asn Ala Ser Ile Val Cys Thr Val
                160                 165                 170
GTG AAC AGG ACA GTA GGT GTT TGG AGA CCA CAC CCT CCT GCC TGT CAA     736
Val Asn Arg Thr Val Gly Val Trp Arg Pro His Pro Pro Ala Cys Gln
            175                 180                 185
AAA ATC GTC TGC CAT CGG CCG CAG ATT CCG AAG GGA TAC TTG GCC CCT     784
Lys Ile Val Cys His Arg Pro Gln Ile Pro Lys Gly Tyr Leu Ala Pro
        190                 195                 200
GGA TTT CGA CAG TTC TAT GCG TAC AGA GAC GCT CTT GAG ATC CGA TGC     832
Gly Phe Arg Gln Phe Tyr Ala Tyr Arg Asp Ala Leu Glu Ile Arg Cys
    205                 210                 215
AAG AAG GGT TTT ATC CTC AGA GGC AGC AGT GTG ATC CAC TGT GAA GCA     880
Lys Lys Gly Phe Ile Leu Arg Gly Ser Ser Val Ile His Cys Glu Ala
220                 225                 230                 235
AAT GGC GAG TGG TTT CCT TCT ATC CCC ACC TGT GAA CCC AAT GGT TGT     928
Asn Gly Glu Trp Phe Pro Ser Ile Pro Thr Cys Glu Pro Asn Gly Cys
                240                 245                 250
ACC AAT ATA CCA GAT ATT TCC TAT GCT TCC TGG GAG GGA TAT AAG TTT     976
Thr Asn Ile Pro Asp Ile Ser Tyr Ala Ser Trp Glu Gly Tyr Lys Phe
            255                 260                 265
CCA TTA AGA AAT TTT GAA GTA TTT GAA ATT GGG GCC AAA TTG AAA TAC    1024
Pro Leu Arg Asn Phe Glu Val Phe Glu Ile Gly Ala Lys Leu Lys Tyr
        270                 275                 280
CAG TGC AAG CCT GGT TAT CGA GCA AGT CTT AAC GAT CCC CAG ACT GTG    1072
Gln Cys Lys Pro Gly Tyr Arg Ala Ser Leu Asn Asp Pro Gln Thr Val
    285                 290                 295
ACT TGT CAG GAA AAT CTG ACT TGG TCA TCT ACT AAT GGA TGT GAA AGG    1120
Thr Cys Gln Glu Asn Leu Thr Trp Ser Ser Thr Asn Gly Cys Glu Arg
300                 305                 310                 315
```

```
ATA TGT TGC CCA ACA CCA GAT ATG GAG AAA ATC AAA ATT GTG AGT GAA   1168
Ile Cys Cys Pro Thr Pro Asp Met Glu Lys Ile Lys Ile Val Ser Glu
                320                 325                 330
AGG AGA GAT TTC ACT GGC ACA TGC ATC TAT GCC TAT GGA GAC TAT GTT   1216
Arg Arg Asp Phe Thr Gly Thr Cys Ile Tyr Ala Tyr Gly Asp Tyr Val
            335                 340                 345
TTC TAC ATT TGT AAT GAA GGC TCT TAC CCT ATG TCT ACG GAT GGA AGG   1264
Phe Tyr Ile Cys Asn Glu Gly Ser Tyr Pro Met Ser Thr Asp Gly Arg
        350                 355                 360
AGT TCA TGT CAA GCA GAT GGA AAG TGG GAC CCT GCA ATA CCA TCA TGT   1312
Ser Ser Cys Gln Ala Asp Gly Lys Trp Asp Pro Ala Ile Pro Ser Cys
    365                 370                 375
CAG GCA GAC TCA GGC CTG CAA AAC CGT CTT GCT CTT TTC ACC TTC CCA   1360
Gln Ala Asp Ser Gly Leu Gln Asn Arg Leu Ala Leu Phe Thr Phe Pro
380                 385                 390                 395
AAC ATA TCA GAA ACC AAT GTG ACA AAC AAA ACC TAT CTA TTT GGT CAT   1408
Asn Ile Ser Glu Thr Asn Val Thr Asn Lys Thr Tyr Leu Phe Gly His
                400                 405                 410
GAA GAA AAC TCA ACC GAG CAT GCC ATG AAA GGT GTG TGT CTC AAA CCA   1456
Glu Glu Asn Ser Thr Glu His Ala Met Lys Gly Val Cys Leu Lys Pro
            415                 420                 425
ATG GTC ATA AAT GGA AAC CTG TCT GTG GAG AGA GTT ATC TAT GCT GAA   1504
Met Val Ile Asn Gly Asn Leu Ser Val Glu Arg Val Ile Tyr Ala Glu
        430                 435                 440
CTG GAA AAT ATC ACC ATT CAA TGT GAT CCT GGA TAT ACC ATA GTT GGT   1552
Leu Glu Asn Ile Thr Ile Gln Cys Asp Pro Gly Tyr Thr Ile Val Gly
    445                 450                 455
TCA CCA AAC ATC ATT TGT TCA AAC AGA ACG TGG TAC CCT GAG GTA CCC   1600
Ser Pro Asn Ile Ile Cys Ser Asn Arg Thr Trp Tyr Pro Glu Val Pro
460                 465                 470                 475
AGC TGT CAG ATG GAG GTC CTA GAA GAC TGC AGA ATA GTG AGC AGA GGC   1648
Ser Cys Gln Met Glu Val Leu Glu Asp Cys Arg Ile Val Ser Arg Gly
                480                 485                 490
GCA CAA CTC TTG CAT TGC CTC TCA AGC CCA GAA GAT GTG CAC AGG GCC   1696
Ala Gln Leu Leu His Cys Leu Ser Ser Pro Glu Asp Val His Arg Ala
            495                 500                 505
CTG AAG GTG TAC AAG CTG TTT CTA GAG ATC GAA CGA TTG GAA CAT CAG   1744
Leu Lys Val Tyr Lys Leu Phe Leu Glu Ile Glu Arg Leu Glu His Gln
        510                 515                 520
```

```
AAA GAG AAG TGG ATA CAG TTA CAC AGG AAA CCT CAG TCT ATG AAA ATT   1792
Lys Glu Lys Trp Ile Gln Leu His Arg Lys Pro Gln Ser Met Lys Ile
    525                 530                 535
AAT AGG TCA TTT AGA CTT TGC AAT TAGTGGACAG AGTCTCCACG CAGCAGCTGG  1846
Asn Arg Ser Phe Arg Leu Cys Asn
540                 545
CCTGACACAC ACCACCACTC CTGTTAAAAG GTTTTGGGGC CAGAAGAACT TAGGTTGTGT 1906
TTTTTTTTTT TCTGTTTTGT TTTTGAAACT ATGTCATTCA TTAGTGTCTA GCCATTGTGG 1966
TTCAATTTGT CTACTTGGCT GGACTCACTG AAATAGTTAA ATATAAACTG TCAAATGTCA 2026
```

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 579 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

```
Met Ile Thr Trp Ser Phe Ile Asp Leu Trp Arg Thr Ser His Ser Thr
-32      -30                 -25                 -20
Leu Phe Gln Met Thr Leu Ala Thr Val Leu Met Ala Pro Val Leu Gly
    -15                 -10                 -5
Asp Cys Gly Pro Pro Pro Leu Leu Pro Phe Ala Ser Pro Thr Asn Gln
  1               5                  10                  15
Leu Tyr Glu Ser Thr Thr Phe Pro Ser Gly Thr Val Leu Lys Tyr Thr
            20                  25                  30
Cys His His Gly Phe Lys Arg Val Asn Ser Ser His Leu Ser Cys Asp
        35                  40                  45
Glu Asn Gly Ser Trp Val Tyr Ser Thr Phe Cys Ala Arg Lys Arg Cys
    50                  55                  60
Lys Asn Pro Gly Glu Leu Val Asn Gly Lys Val Glu Ile Pro Ser Asp
65                  70                  75                  80
Leu Leu Val Gly Ser Ile Ile Glu Phe Ser Cys Ser Lys Gly Tyr Leu
                85                  90                  95
Leu Ile Gly Ser Ala Thr Ser Arg Cys Glu Val Gln Gly Lys Gly Val
            100                 105                 110
Asp Trp Ser Asp Ser Leu Pro Glu Cys Val Ile Ala Thr Cys Glu Pro
        115                 120                 125
```

Pro Pro Pro Ile Ser Asn Gly Lys His Ser Gly Arg Asp Asp Asp Leu
130 135 140
Tyr Thr Phe Gly Ser Val Val Ile Tyr Asn Cys Asp Pro Thr Phe Thr
145 150 155 160
Leu Leu Gly Asn Ala Ser Ile Val Cys Thr Val Val Asn Arg Thr Val
165 170 175
Gly Val Trp Arg Pro His Pro Pro Ala Cys Gln Lys Ile Val Cys His
180 185 190
Arg Pro Gln Ile Pro Lys Gly Tyr Leu Ala Pro Gly Phe Arg Gln Phe
195 200 205
Tyr Ala Tyr Arg Asp Ala Leu Glu Ile Arg Cys Lys Lys Gly Phe Ile
210 215 220
Leu Arg Gly Ser Ser Val Ile His Cys Glu Ala Asn Gly Glu Trp Phe
225 230 235 240
Pro Ser Ile Pro Thr Cys Glu Pro Asn Gly Cys Thr Asn Ile Pro Asp
245 250 255
Ile Ser Tyr Ala Ser Trp Glu Gly Tyr Lys Phe Pro Leu Arg Asn Phe
260 265 270
Glu Val Phe Glu Ile Gly Ala Lys Leu Lys Tyr Gln Cys Lys Pro Gly
275 280 285
Tyr Arg Ala Ser Leu Asn Asp Pro Gln Thr Val Thr Cys Gln Glu Asn
290 295 300
Leu Thr Trp Ser Ser Thr Asn Gly Cys Glu Arg Ile Cys Cys Pro Thr
305 310 315 320
Pro Asp Met Glu Lys Ile Lys Ile Val Ser Glu Arg Arg Asp Phe Thr
325 330 335
Gly Thr Cys Ile Tyr Ala Tyr Gly Asp Tyr Val Phe Tyr Ile Cys Asn
340 345 350
Glu Gly Ser Tyr Pro Met Ser Thr Asp Gly Arg Ser Ser Cys Gln Ala
355 360 365
Asp Gly Lys Trp Asp Pro Ala Ile Pro Ser Cys Gln Ala Asp Ser Gly
370 375 380
Leu Gln Asn Arg Leu Ala Leu Phe Thr Phe Pro Asn Ile Ser Glu Thr
385 390 395 400
Asn Val Thr Asn Lys Thr Tyr Leu Phe Gly His Glu Glu Asn Ser Thr
405 410 415
Glu His Ala Met Lys Gly Val Cys Leu Lys Pro Met Val Ile Asn Gly
420 425 430
Asn Leu Ser Val Glu Arg Val Ile Tyr Ala Glu Leu Glu Asn Ile Thr
435 440 445

Ile Gln Cys Asp Pro Gly Tyr Thr Ile Val Gly Ser Pro Asn Ile Ile
450 455 460
Cys Ser Asn Arg Thr Trp Tyr Pro Glu Val Pro Ser Cys Gln Met Glu
465 470 475 480
Val Leu Glu Asp Cys Arg Ile Val Ser Arg Gly Ala Gln Leu Leu His
485 490 495
Cys Leu Ser Ser Pro Glu Asp Val His Arg Ala Leu Lys Val Tyr Lys
500 505 510
Leu Phe Leu Glu Ile Glu Arg Leu Glu His Gln Lys Glu Lys Trp Ile
515 520 525
Gln Leu His Arg Lys Pro Gln Ser Met Lys Ile Asn Arg Ser Phe Arg
530 535 540
Leu Cys Asn
545

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 23 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

CCNTTYGCNW SNCCNACNAA YCA 23

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 23 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

TARAANACRT ARTCNCCRTA NGC 23

(2) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 23 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

GCNWSNCCNA CNAAYCARYT NTA 23

(2) INFORMATION FOR SEQ ID NO:6:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 17 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

ACCTCGCATC ACGGCTT 17

(2) INFORMATION FOR SEQ ID NO:7:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 18 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

GAAGGAAACC ACTCGCCA 18

(2) INFORMATION FOR SEQ ID NO:8:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 89 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

ACCAATGTGA CGAACAAGAC CTACTTATTC GGTCATGAAG AAAATTCTAC TGAGCATGCC 60

ATGAAAGGTG TGGGTCCTGC AACTTGAGG 89

(2) INFORMATION FOR SEQ ID NO:9:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 99 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

GGTCTTGTTC GTCACATTGG TTTCAGATAT ATTTGGGAAA GTGAATAAAG GCTAATCTAT 60

TTTGTAAACC AGAGTCTGCT TGAGAGGTCC AAACCAAAC 99

(2) INFORMATION FOR SEQ ID NO:10:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 23 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

TAATACGACT CACTATAGGG AGA 23

(2) INFORMATION FOR SEQ ID NO:11:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 20 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

AATAACCCTC ACTAAAGGGA 20

(2) INFORMATION FOR SEQ ID NO:12:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1956 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

AAGCTTGCAT GCCTGCAGGT CAACATGGTG GAGCACGACA CTCTCGTCTA CTCCAAGAAT 60
ATCAAAGATA CAGTCTCAGA AGACCAGAGG GCTATTGAGA CTTTTCAACA AAGGGTAATA 120
TCGGGAAACC TCCTCGGATT CCATTGCCCA GCTATCTGTC ACTTCATCGA AAGGACAGTA 180
GAAAAGGAAG ATGGCTTCTA CAAATGCCAT CATTGCGATA AAGGAAAGGC TATCGTTCAA 240
GAATGCCTCT ACCGACAGTG GTCCCAAAGA TGGACCCCCA CCCACGAGGA ACATCGTGGA 300
AAAAGAAGAC GTTCCAACCA CGTCTTCAAA GCAAGTGGAT TGATGTGATA TCATGGTGGA 360
GCACGACACT CTCGTCTACT CCAAGAATAT CAAAGATACA GTCTCAGAAG ACCAGAGGGC 420
TATTGAGACT TTTCAACAAA GGGTAATATC GGGAAACCTC CTCGGATTCC ATTGCCCAGC 480
TATCTGTCAC TTCATCGAAA GGACAGTAGA AAAGGAAGAT GGCTTCTACA AATGCCATCA 540
TTGCGATAAA GGAAAGGCTA TCGTTCAAGA ATGCCTCTAC CGACAGTGGT CCCAAAGATG 600
GACCCCCACC CACGAGGAAC ATCGTGGAAA AAGAAGACGT TCCAACCACG TCTTCAAAGC 660
AAGTGGATTG ATGTGATATC TCCACTGACG TAAGGGATGA CGCACAATCC CACTATCCTT 720
CGCAAGACCC TTCCTCTATA TAAGGAAGTT CATTTCATTT GGAGAGGACC TCGAGAATTC 780
TCAACACAAC ATATACAAAA CAAACGAATC TCAAGCAATC AAGCATTCTA CTTCTATTGC 840

AGCAATTTAA ATCATTTCTT TTAAAGCAAA AGCAATTTTC TGAAAATTTT CACCATTTAC 900
GAACGATAGC CATGGGCTGG AGCTGGATCT TTCTCTTCCT CCTGTCAGGA GCTGCAGGTG 960
GGTACCTTAA TGATGATAAG TTATATCGGG CAGATTCTAG ACCTCCTGAT GAAATAAAGC 1020
AGTCAGGTGG TCTTATGCCA AGAGGACAGA GTGAGTACTT TGACCGAGGT ACTCAAATGA 1080
ATATCAACCT TTATGATCAT GCAAGAGGAA CTCAGACGGG ATTTGTTAGG CACGATGATG 1140
GATATGTTTC CACCTCAATT AGTTTGAGAA GTGCCCACTT AGTGGGTCAA ACTATATTGT 1200
CTGGTCATTC TACTTATTAT ATATATGTTA TAGCCACTGC ACCCAACATG TTTAACGTTA 1260
ATGATGTATT AGGGGCATAC AGTCCTCATC CAGATGAACA AGAAGTTTCT GCTTTAGGTG 1320
GGATTCCATA CTCCCAAATA TATGGATGGT ATCGAGTTCA TTTTGGGGTG CTTGATGAAC 1380
AATTACATCG TAATAGGGGC TACAGAGATA GATATTACAG TAACTTAGAT ATTGCTCCAG 1440
CAGCAGATGG TTATGGATTG GCAGGTTTCC CTCCGGAGCA TAGAGCTTGG AGGGAAGAGC 1500
CGTGGATTCA TCATGCACCG CCGGGTTGTG GGAATGCTCC AAGATCATCG ATGAGTAATA 1560
CTTGCGATGA AAAAACCCAA AGTCTAGGTG TAAAATTCCT TGACGAATAC CAATCTAAAG 1620
TTAAAAGACA AATATTTTCA GGCTATCAAT CTGATATTGA TACACATAAT AGAATTAAGG 1680
ATGAATTATG AGGTACCCGG GGATCTGAAA AGGATGAACT TTAGTCTAGA GTCCGCAAAT 1740
CACCAGTCTC TCTCTACAAA TCTATCTCTC TCTATTTTCT CCAGAATAAT GTGTGAGTAG 1800
TTCCCAGATA AGGGAATTAG GGTTCTTATA GGGTTTCGCT CATGTGTTGA GCATATAAGA 1860
AACCCTTAGT ATGTATTTGT ATTTGTAAAA TACTTCTATC AATAAAATTT CTAATTCCTA 1920
AAACCAAAAT CCAGTGACCT GCAGGCATGC AAGCTT 1956

(2) INFORMATION FOR SEQ ID NO:13:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1525 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

AAGCTTGCAT GCCTGCAGGT CAACATGGTG GAGCACGACA CTCTCGTCTA CTCCAAGAAT 60
ATCAAAGATA CAGTCTCAGA AGACCAGAGG GCTATTGAGA CTTTTCAACA AAGGGTAATA 120
TCGGGAAACC TCCTCGGATT CCATTGCCCA GCTATCTGTC ACTTCATCGA AAGGACAGTA 180
GAAAAGGAAG ATGGCTTCTA CAAATGCCAT CATTGCGATA AAGGAAAGGC TATCGTTCAA 240
GAATGCCTCT ACCGACAGTG GTCCCAAAGA TGGACCCCCA CCCACGAGGA ACATCGTGGA 300
AAAAGAAGAC GTTCCAACCA CGTCTTCAAA GCAAGTGGAT TGATGTGATA TCATGGTGGA 360
GCACGACACT CTCGTCTACT CCAAGAATAT CAAAGATACA GTCTCAGAAG ACCAGAGGGC 420
TATTGAGACT TTTCAACAAA GGGTAATATC GGGAAACCTC CTCGGATTCC ATTGCCCAGC 480
TATCTGTCAC TTCATCGAAA GGACAGTAGA AAAGGAAGAT GGCTTCTACA AATGCCATCA 540

TTGCGATAAA GGAAAGGCTA TCGTTCAAGA ATGCCTCTAC CGACAGTGGT CCCAAAGATG 600
GACCCCCACC CACGAGGAAC ATCGTGGAAA AAGAAGACGT TCCAACCACG TCTTCAAAGC 660
AAGTGGATTG ATGTGATATC TCCACTGACG TAAGGGATGA CGCACAATCC CACTATCCTT 720
CGCAAGACCC TTCCTCTATA TAAGGAAGTT CATTTCATTT GGAGAGGACC TCGAGAATTC 780
TCAACACAAC ATATACAAAA CAAACGAATC TCAAGCAATC AAGCATTCTA CTTCTATTGC 840
AGCAATTTAA ATCATTTCTT TTAAAGCAAA AGCAATTTTC TGAAAATTTT CACCATTTAC 900
GAACGATAGC CATGGGTTGG TCTTGGATTT TTCTTTTGTC AGGAGCTGCA GGAGGGTACC 960
AACTCCTCAA AACATCACAG ATCTTTGTGC TGAATCACAT AACACTCAAA TATATACCTT 1020
GAACGATAAG ATTTTTAGTT ACACCGAGAG CTTAGCCGGT AAAAGGGAAA TGGCTATTAT 1080
AACTTTTAAG AACGGTGCTA TTTTCCAGGT CGAAGTTCCA TCCTCTCAAC ATATAGACAG 1140
TCAAAAGAAG GTATCGAAAG GATGAAGGAT ACACTCAGAA TTGCATACCT TACAGAGGCT 1200
AAAGTCGAAA AGTTGTGTGT ATGGAATAAC AAAACCCCAC ATGCTATTGC TGCCATTAGT 1260
ATGGCAAACG GATCTGAAAA GGATGAACTT TAGTCTAGAG TCCGCAAATC ACCAGTCTCT 1320
CTCTACAAAT CTATCTCTCT CTATTTTCTC CAGAATAATG TGTGAGTAGT TCCCAGATAA 1380
GGGAATTAGG GTTCTTATAG GGTTTCGCTC ATGTGTTGAG CATATAAGAA ACCCTTAGTA 1440
TGTATTTGTA TTTGTAAAAT ACTTCTATCA ATAAAATTTC TAATTCCTAA AACCAAAATC 1500
CAGTGACCTG CAGGCATGCA AGCTT 1525

(2) INFORMATION FOR SEQ ID NO:14:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 17 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

Cys Asn Ile Ser Glu Thr Asn Val Thr Asn Lys Thr Tyr Leu Phe Gly
1 5 10 15

His

(2) INFORMATION FOR SEQ ID NO:15:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 17 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:

Cys His Gly Phe Leu Tyr Thr Lys Asn Thr Val Asn Thr Glu Ser Ile
1 5 10 15

Asn

(2) INFORMATION FOR SEQ ID NO:16:
(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 60 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear
(ii) MOLECULE TYPE: cDNA
(iii) HYPOTHETICAL: NO
(iv) ANTI-SENSE: NO
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:16:

TCTGAATAAG ACGTAGGAAG ACTCTAGTGT TGCCTTGGGT TTTTCGCCCC TCCTCGTGCC 60

(2) INFORMATION FOR SEQ ID NO:17:
(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 60 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear
(ii) MOLECULE TYPE: cDNA
(iii) HYPOTHETICAL: NO
(iv) ANTI-SENSE: NO
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:17:

GTTGTTGACT CATACAACTG GTTGGTTGGA GAAGCAAACG GTAAAAGGGG TGGAGGTCCA 60

(2) INFORMATION FOR SEQ ID NO:18:
(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 60 base pairs
(B) TYPE: nucleic acid

(C) STRANDEDNESS: single
(D) TOPOLOGY: linear
(ii) MOLECULE TYPE: cDNA
(iii) HYPOTHETICAL: NO
(iv) ANTI-SENSE: NO
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:18:

CCCTTCTTGC ATCGGATCTC AAGAGCGTCT CTGTACGCAT AGAACTGTCG AAATCCAGGG 60

(2) INFORMATION FOR SEQ ID NO:19:
(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 60 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear
(ii) MOLECULE TYPE: cDNA
(iii) HYPOTHETICAL: NO
(iv) ANTI-SENSE: NO
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:19:

CCTTTCATGG CATGCTCGGT TGAGTTTTCT TCATGACCAA ATAGATAGGT TTTGTTTGTC 60

What Is Claimed Is:

1. An immunocontraceptive vaccine composition comprising an antigenic polypeptide having an sp56 antigenic determinant in an amount effective to induce anti-sp56 antibodies in a mammalian subject in combination with a pharmaceutically acceptable excipient.

2. The composition of claim 1 wherein said antigenic determinant includes an sp56 species specific sequence (sss).

3. The composition of claim 1 wherein said antigenic determinant includes a sp56 Sushi domain.

4. The composition of claim 1 wherein said antigenic polypeptide is selected from the group consisting of a synthetic polypeptide, a hybrid molecule, a fusion protein, a recombinant protein, a chimeric polypeptide molecule and an isolated protein.

5. The composition of claim 1 wherein said mammalian subject is selected from the group consisting of rodent, ungulate, marsupial, primate, porcine, canine, feline and human.

6. The composition of claim 1 wherein said antigenic polypeptide comprises a polypeptide selected from the group consisting of SEQ ID NOs 2:-32 to 547, 2:1-547, 2:1-477, 2:1-59, 2:3-32, 2:34-46, 2:64-121, 2:65-90, 2:92-104, 2:126-186, 2:127-154, 2:156-168, 2:191-246, 2:192-218, 2:220-232, 2:251-313, 2:252-284, 2:286-300, 2:317-379, 2:319-350, 2:352-365, 2:319-338, 2:340-350, 2:424-477, 2:425-450, 2:452-464, 2:380-423, 2:383-398, 2:396-411 and 2:406-421.

7. The composition of claim 6 wherein said composition comprises two or more of said antigenic polypeptides.

8. The composition of claim 1 further comprising a carrier, an immune stimulator, an antigenic molecule or an adjuvant.

9. The composition of claim 8 wherein said carrier is a polypeptide selected from the group consisting of cholera toxin subunit A, cholera toxin subunit B, diphtheria toxin, influenza virus HA protein, murine leukemia virus coat protein, salmonella surface protein, and tobacco mosaic virus coat protein.

10. The composition of claim 1 further comprising a stabilizer.

11. The composition of claim 8 wherein said carrier is a non-toxic, edible carrier.

12. The composition of claim 11 wherein said carrier is plant material.

13. The composition of claim 11 wherein said composition is formulated as a bait or feed.

14. The composition of claim 4 wherein said hybrid molecule comprises a first polypeptide domain consisting of the sp56 antigenic determinant operatively linked a second polypeptide domain.

15. The composition of claim 14 wherein said second polypeptide domain is selected from the group consisting of cholera toxin subunit A, cholera toxin subunit B, diphtheria toxin, influenza virus HA protein, murine leukemia virus coat protein, salmonella surface protein, and tobacco mosaic virus coat protein.

16. The composition of claim 14 wherein said hybrid molecule comprises a polypeptide having a cholera toxin subunit A polypeptide domain operatively linked to a polypeptide defining a species specific sequence (sss) of sp56.

17. The composition of claim 16 werein said sss polypeptide is selected from the group consisting of 2:380-423, 2:383-398, 2:396-411 and 2:406-421.

18. The composition of claim 14 wherein said hybrid molecule is a fusion protein and said operative linkage is a peptide bond between said first and second polypeptide domains.

19. The composition of claim 14 wherein said hybrid molecule is a chimeric polypeptide molecule and said operative linkage is a chemical cross-link.

20. An immunocontraceptive method comprising administering to a mammalian subject an immunocontraceptive vaccine composition according to claim 1 in an amount effective to induce production of anti-sp56 antibodies.

21. The method of claim 20 wherein said administering route is selected from the group consisting of oral, intravenous, subcutaneous, intramuscular and intraperitoneal.

22. The method of claim 21 wherein said oral administration is by ingestion.

23. The method of claim 21 wherein said mammal is selected from the group consisting of selected from the group consisting of rodent, ungulate, marsupial, primate, porcine, canine, feline and human.

24. A method for determining infertility in a mammal comprising the steps of:

a) admixing a body sample containing antibodies from said mammal with an antigenic polypeptide having an sp56 antigenic determinant of said mammal;

b) maintaining said admixture under conditions sufficient for said polypeptide to immunoreact with any of said antibodies to form an immunoreaction complex; and

c) detecting the presence of said immunoreaction complex in said admixture.

25. An antibody that immunoreacts with an antigenic polypeptide having an sp56 antigenic determinant, wherein said polypeptide is no more than about 65 amino acid residues in length and comprises a sequence selected from the group consisting of 2:1-59, 2:3-32, 2:34-46, 2:64-121, 2:65-90, 2:92-104, 2:126-186, 2:127-154, 2:156-168, 2:191-246, 2:192-218, 2:220-232, 2:251-313, 2:252-284, 2:286-300, 2:317-379, 2:319-350, 2:352-365, 2:319-338, 2:340-350, 2:424-477, 2:425-450, 2:452-464, 2:380-423, 2:383-398, 2:396-411 and 2:406-421.

26. The antibody of claim 25 wherein said polypeptide consists of a amino acid residue sequence selected from the group consisting of 2:1-59, 2:3-32, 2:34-46, 2:64-121, 2:65-90, 2:92-104, 2:126-186, 2:127-154, 2:156-168, 2:191-246, 2:192-218, 2:220-232, 2:251-313, 2:252-284, 2:286-300, 2:317-379, 2:319-350, 2:352-365, 2:319-338, 2:340-350, 2:424-477, 2:425-450, 2:452-464, 2:380-423, 2:383-398, 2:396-411 and 2:406-421.

27. The antibody of claim 25 wherein said antibody is a monoclonal antibody.

28. A method for determining male fertility in a mammal comprising the steps of:

a) admixing a body sample containing sperm from said mammal with an antibody that immunoreacts with an antigenic polypeptide having an sp56 antigenic determinant of said mammal;

b) maintaining said admixture under conditions sufficient for said antibody to immunoreact with any of said sperm containing said antigenic determinant in said body sample to form an immunoreaction complex; and

c) detecting the presence of said immunoreaction complex in said admixture.

FIGURE 1

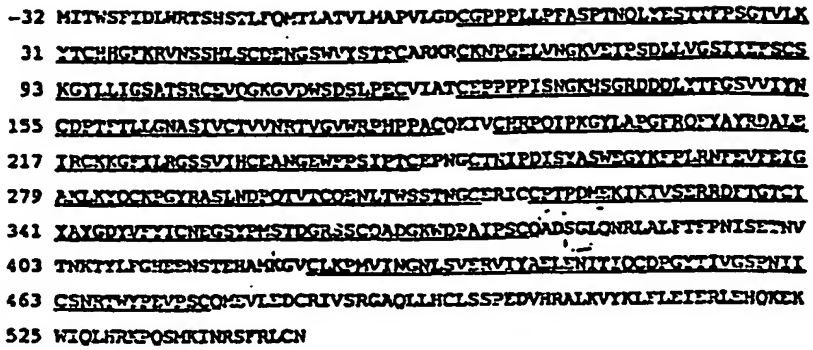
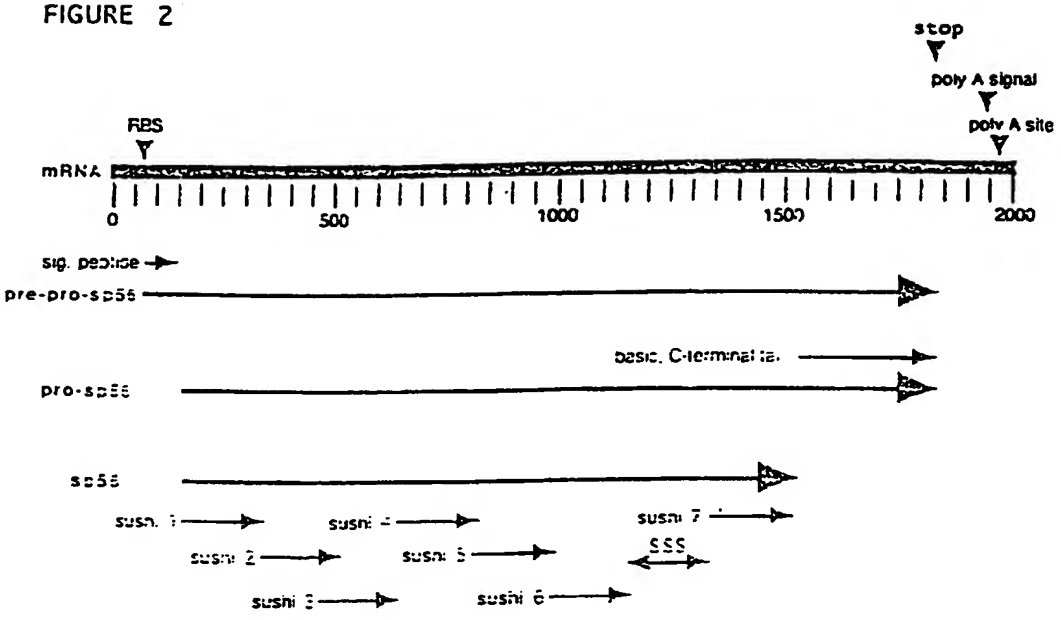

-32 MITWSFIDLWRTSHSTLFQMTLATVLMAPVLGDCGPPPLLPFASPTNQLYESTTFPSGTVLK
31 YTCHRGFKRVNSSHLSCDENGSWVYSTFCARKRCKNPGELVNGKVETPSDLLVGSIIEYSCS
93 KGYLLIGSATSRCEVQGKGVDWSDSLPECVIATCEPPPPISNGKHSGRDDDLYTFGSVVIYN
155 CDPTFTLLGNASIVCTVVNRTVGVWRPHPPACQKIVCHRPQIPKGYLAPGFRQFYAYRDALE
217 IRCNKGFILRGSSVIHCEANGEWFPSIPTCEPNGGLTNIPDISYASWEGYKFPLRNFEVTEIG
279 AYLKYQCKPGYRASLNDPQTVTCQENLTWSSTNGCERICCPTPDMEKIKIVSERRDFNGTGI
341 YAYGDVVFYICNEGSYPMSTDGRSSGQADGKWDPAIPSCQADSGLQNRLALFTFPNISFTNV
403 TNKTYLPGHEENSTEHAMGGVCLKPMVINGNLSVERVIYALFSNITIQCDPGYTIVGSPNII
463 CSNRTWYPEVPSCQMEVLEDCRIVSRGAQLLHCLSSPEDVHRALKVYKLFLEIERLEHQKEK
525 WIQLHRQPQSMKINRSFRLCN

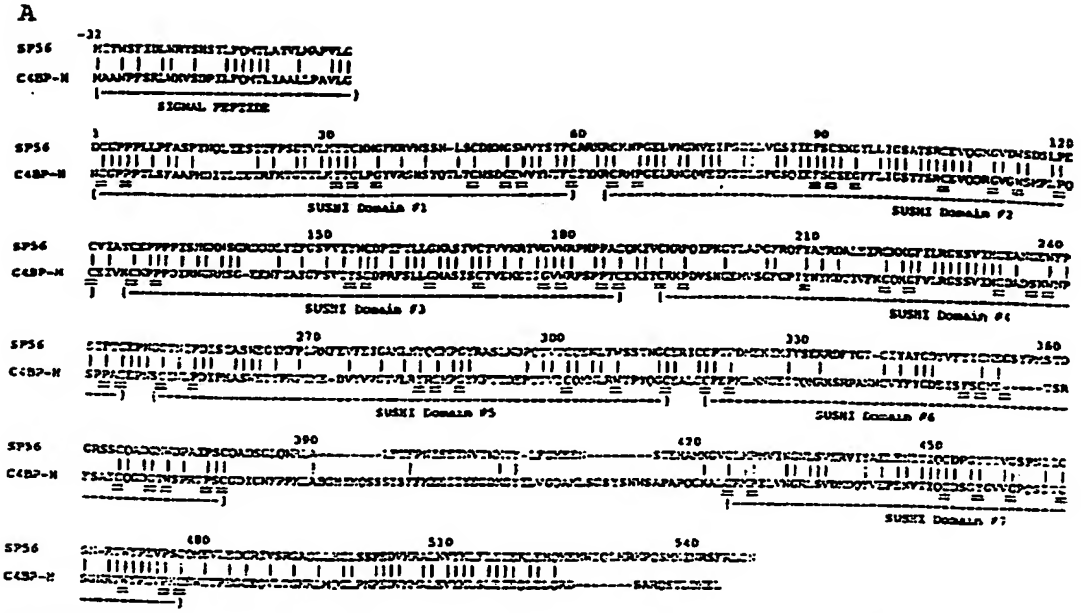
FIGURE 2



Graphic description of the sp56 cDNA, clone 7.1 (GenBank accession number U17108). Base numbers are indicated as "ruler" markings. Ribosome binding site (RBS) and poly A signal (and corresponding experimentally determined poly A site) are indicated on the ruler (mRNA). Polypeptide is indicated by arrows. Positions of Sushi Domains 1 through 7, the signal peptide, the sp56 specific sequence (SSS), and the basic, C-terminal tail of pro-sp56 are indicated.

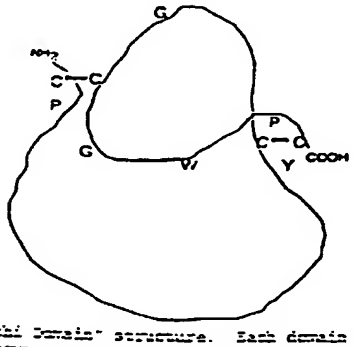
FIGURE 3

A



Translated sp56 cDNA sequence homology with translated human C4 binding protein sequence. Sushi domains are indicated, and virtually invariant amino acids within those sequences are double-underlined. Identical amino acids are indicated by "|". Amino acids are numbered from experimentally determined N-terminal. Signal peptides were identified on the basis of 1) preceding experimentally determined N-termini, 2) following the closest upstream ribosome binding site, 3) being contiguous with the largest open reading frame, and 4) beginning with methionine.

B



Typical "Sushi Domain" structure. Each domain is composed of approximately 60 amino acids, of which 10 (N2- CPYCGGWPC- COOH) are virtually invariant and are located at specific positions in the sequence. The remaining ~50 amino acids evidence considerable divergence. This general structure has the following features: 1) two disulfide bonds (C1-C3 and C2-C4), which form a looped-loop structure, 2) at least one proline near each disulfide bond, providing chain turns necessary for attainment of contiguous "Sushi" domains, and 3) presence of G and W on the inner (hydrophobic) loop.

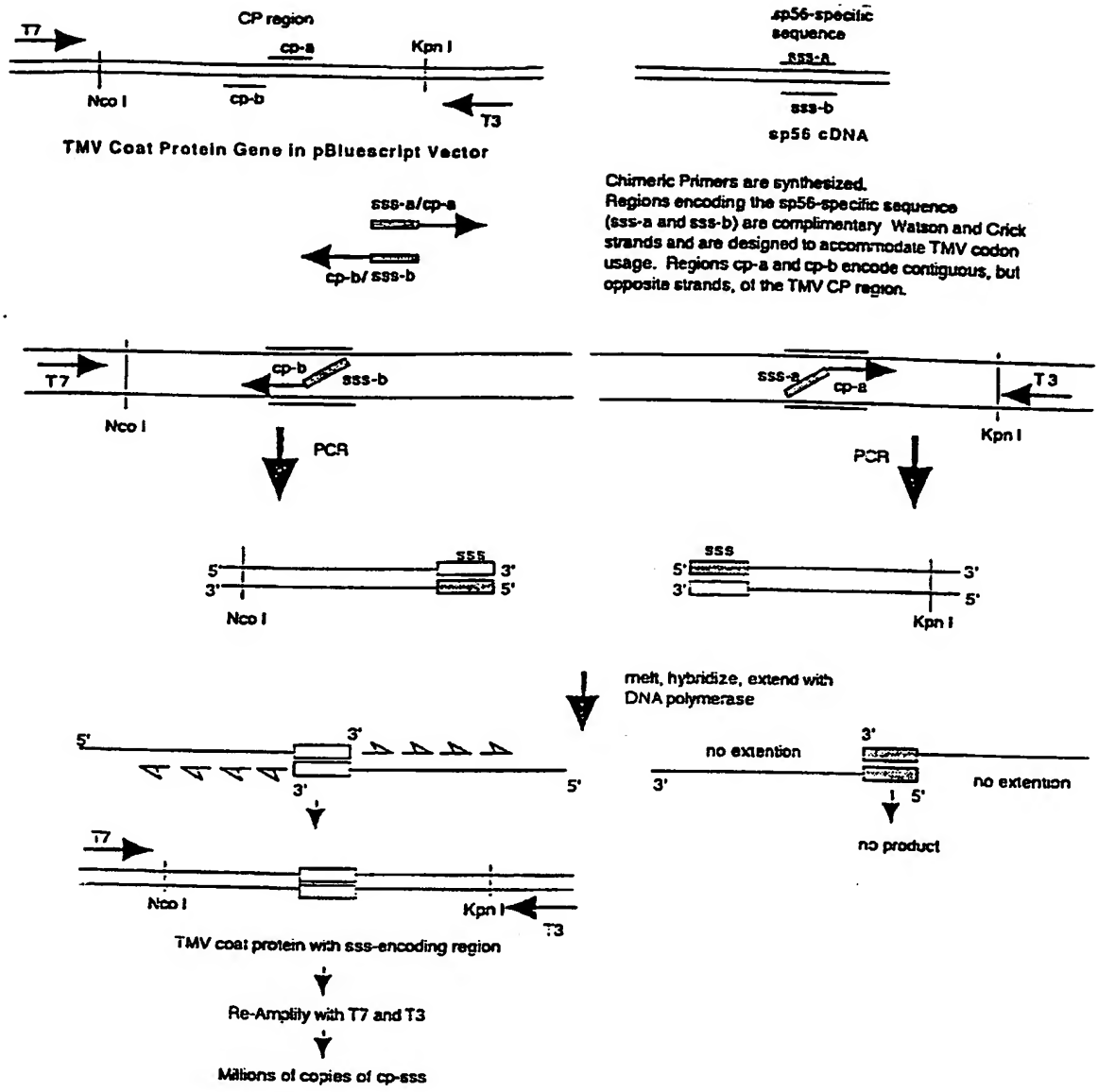

T7
CP region
cp-a
Kpn I
Nco I
cp-b
T3
TMV Coat Protein Gene in pBluescript Vector
sp56-specific sequence
sss-a
sss-b
sp56 cDNA
sss-a/cp-a
cp-b/ sss-b
Chimeric Primers are synthesized. Regions encoding the sp56-specific sequence (sss-a and sss-b) are complimentary Watson and Crick strands and are designed to accommodate TMV codon usage. Regions cp-a and cp-b encode contiguous, but opposite strands, of the TMV CP region.
T7
cp-b
sss-b
Nco I
PCR
sss-a
cp-a
T3
Kpn I
PCR
sss
5'
3'
Nco I
sss
5'
3'
Kpn I
melt, hybridize, extend with DNA polymerase
5'
3'
3'
5'
no extention
3'
3'
5'
no extention
no product
T7
Nco I
Kpn I
T3
TMV coat protein with sss-encoding region
Re-Amplify with T7 and T3
Millions of copies of cp-sss

FIGURE 4

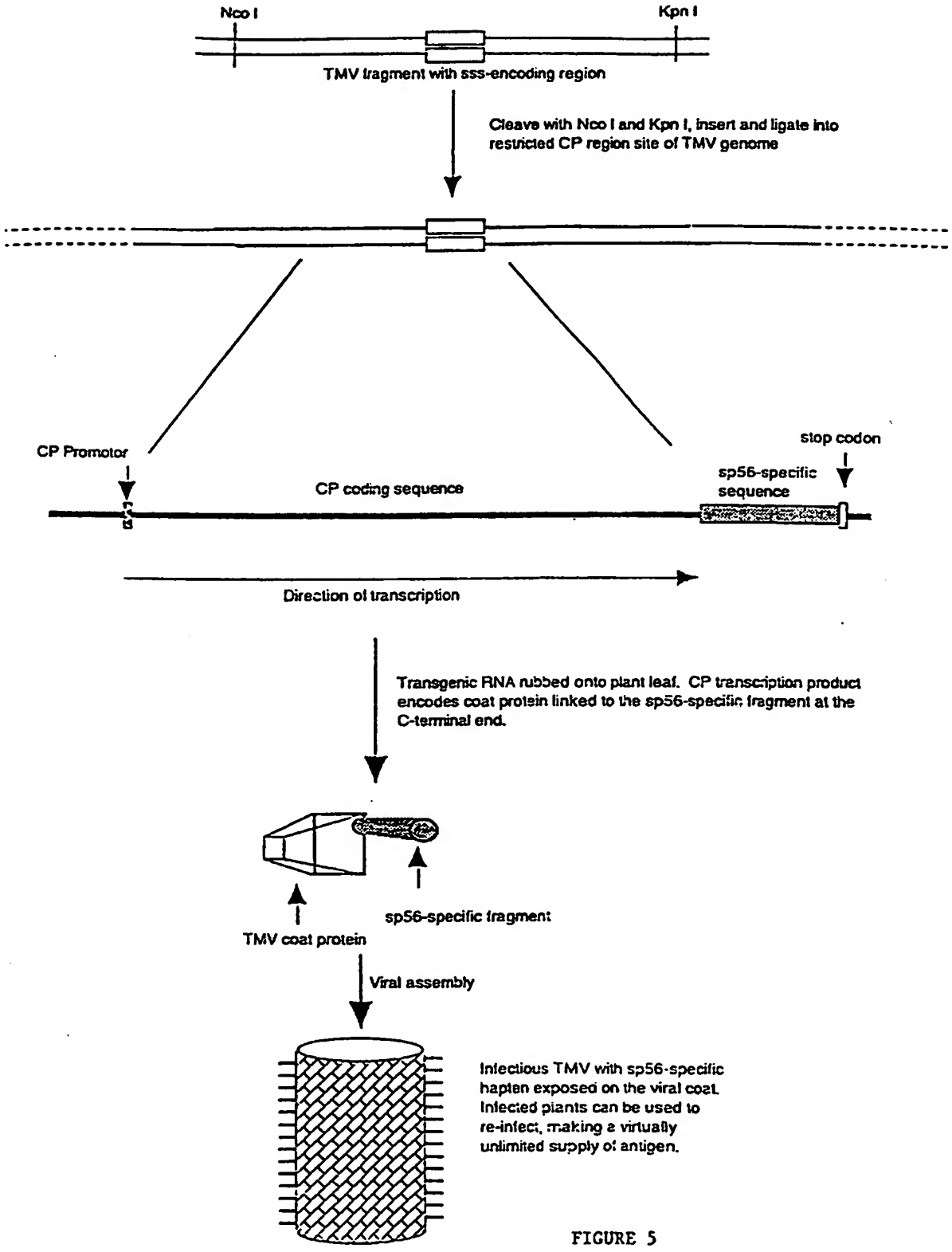

Nco I
Kpn I
TMV fragment with sss-encoding region
Cleave with Nco I and Kpn I, insert and ligate into restricted CP region site of TMV genome
CP Promotor
CP coding sequence
stop codon
sp56-specific sequence
Direction of transcription
Transgenic RNA rubbed onto plant leaf. CP transcription product encodes coat protein linked to the sp56-specific fragment at the C-terminal end.
sp56-specific fragment
TMV coat protein
Viral assembly
Infectious TMV with sp56-specific hapten exposed on the viral coat. Infected plants can be used to re-infect, making a virtually unlimited supply of antigen.

FIGURE 5

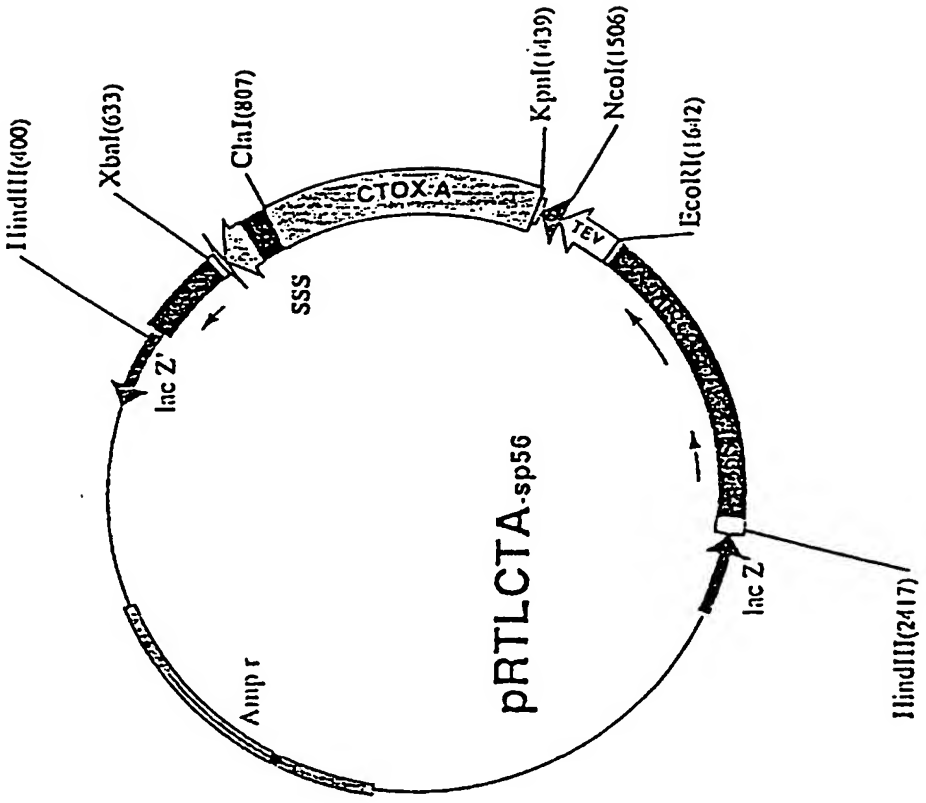

HindIII(400)
XbaI(633)
ClaI(807)
CTOX-A
KpnI(1439)
NcoI(1506)
EcoRI(1642)
TEV
SSS
lac Z'
pRTLCTA-sp56
Amp r
lac Z'
HindIII(2417)

FIGURE 6

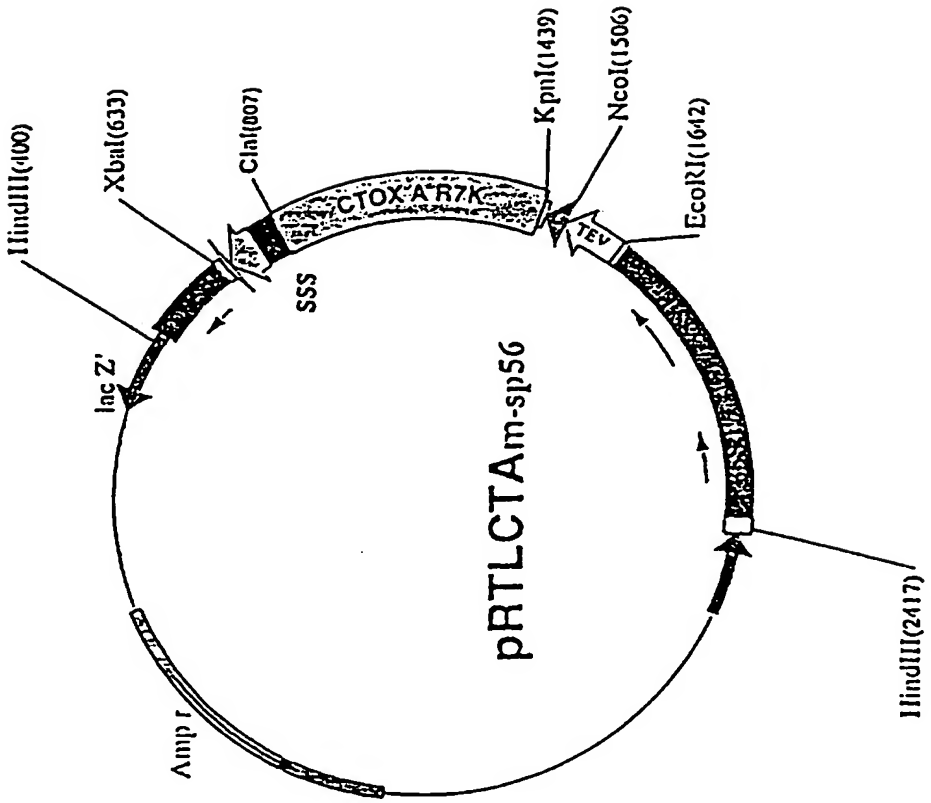

HindIII(100)
XbaI(633)
ClaI(807)
CTOX-A R7K
KpnI(1439)
NcoI(1506)
EcoRI(1642)
TEV
SSS
lac Z'
pRTLCTAm-sp56
Amp r
HindIII(2417)

FIGURE 7

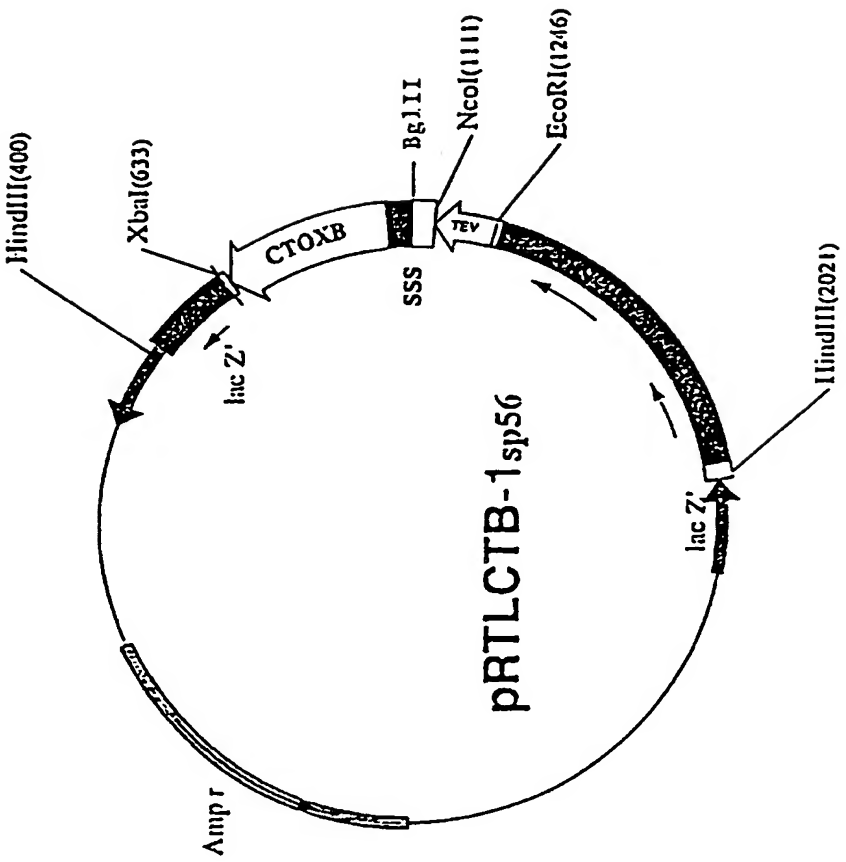

HindIII(400)
XbaI(633)
CTOXB
BglII
NcoI(1111)
EcoRI(1246)
SSS
tev
lac Z'
pRTLCTB-1sp56
HindIII(2021)
lac Z'
Amp r

FIGURE 8

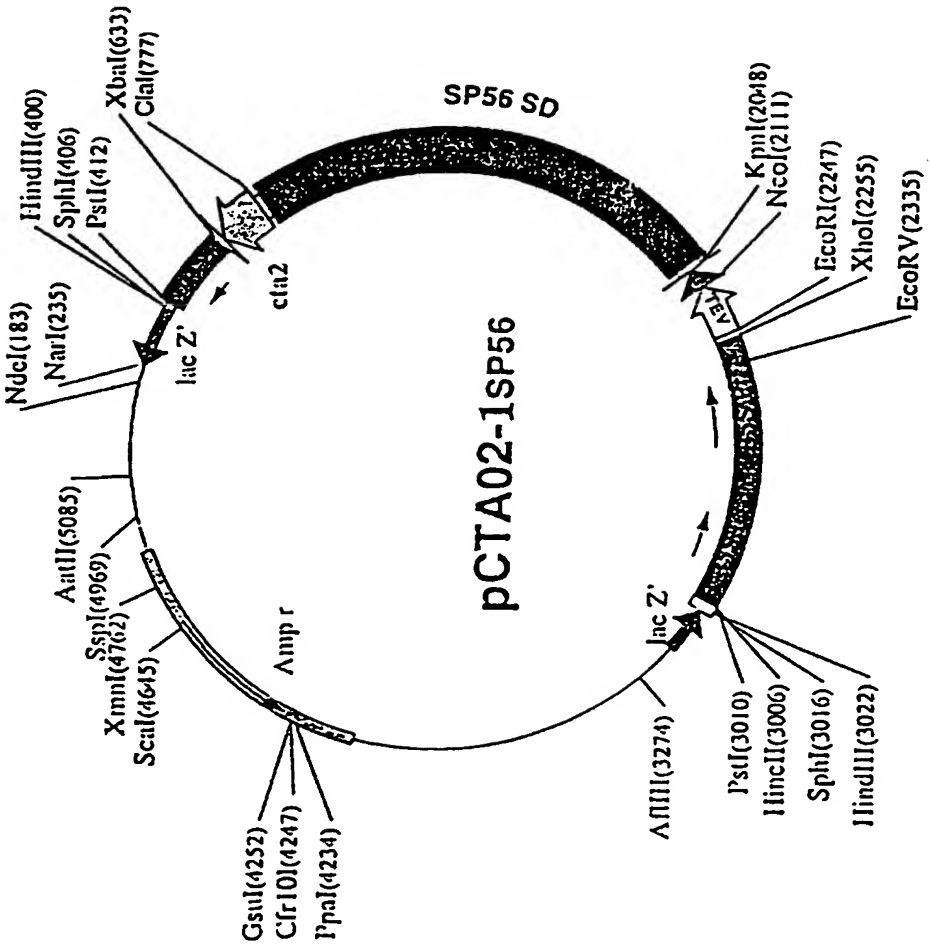
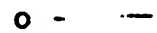
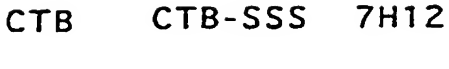

pCTA02-1SP56
SP56 SD
cta2
lac Z'
Amp r
lac Z'
HindIII(400)
SphI(406)
PstI(412)
XbaI(633)
ClaI(777)
NdeI(183)
NarI(235)
AatII(5085)
SspI(4962)
XmnI(4762)
ScaI(4645)
GsuI(4252)
ClrI01(4247)
PpaI(4234)
AflII(3274)
PstI(3010)
HincII(3006)
SphI(3016)
HindIII(3022)
KpnI(2048)
NcoI(2111)
EcoRI(2247)
XhoI(2255)
EcoRV(2335)

FIGURE 9

FIGURE 10





INTERNATIONAL SEARCH REPORT

International application No.
PCT/US97/11719

A. CLASSIFICATION OF SUBJECT MATTER

IPC(6) : C07K 5/00, 14/00; A61K 38/00
US CL : 530/300, 350, 514/2

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

U.S. : 530/300, 350, 514/2

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

APS, DIALOG
search terms: sp56, immunocontraceptive, toxin conjugate

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X -- Y	BLEIL et al. Identification of a ZP3-binding protein on acrosome-intact mouse sperm by photoaffinity cross-linking. Proc. Natl. Acad. Sci. USA. July 1990, Vol. 87, pages 5563-5567. See entire document.	1-5, 7 -------- 6, 8-23
X --- Y	BOOKBINDER et al. Tissue- and species-specific expression of sp56, a mouse sperm fertilization protein. Science. 07 July 1995, Vol. 269, pages 86-89. See entire document.	1-5, 7 -------- 6, 8-23
X -- Y	CHENG et al. Sperm-egg recognition in the mouse: Characterization of sp56, a sperm protein having specific affinity for ZP3. J. Cell Biology. May 1994, Vol. 125, No. 4, pages 867-878. See entire document.	1-5, 7 -------- 6, 8-23

[X] Further documents are listed in the continuation of Box C. [ ] See patent family annex.

*	Special categories of cited documents:	"T"	later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"A"	document defining the general state of the art which is not considered to be of particular relevance	"X"	document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"E"	earlier document published on or after the international filing date	"Y"	document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
"L"	document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)	"&"	document member of the same patent family
"O"	document referring to an oral disclosure, use, exhibition or other means		
"P"	document published prior to the international filing date but later than the priority date claimed		

Date of the actual completion of the international search 26 AUGUST 1997	Date of mailing of the international search report 26 SEP 1997
Name and mailing address of the ISA/US Commissioner of Patents and Trademarks Box PCT Washington, D.C. 20231 Facsimile No. (703) 305-3230	Authorized officer MINH-TAM DAVIS Telephone No. (703) 308-0196

Form PCT/ISA/210 (second sheet)(July 1992)*

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US97/11719

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	THRUSH et al. Immunotoxins: an update. Ann. Rev. Immunol. 1996, Vol. 14, pages 49-71. See reference 45, pages 53, and 65.	9-19
Y	PRIMAKOFF et al. Fully effective contraception in male and female guinea pig immunized with the sperm protein PH-20. Nature. 06 October 1988, Vol. 335, pages 543-546. See entire document.	20-23

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US97/11719

Box I Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)

This international report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☐ Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:

2. ☐ Claims Nos.:
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:

3. ☐ Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

Please See Extra Sheet.

1. ☐ As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.

2. ☐ As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.

3. ☐ As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:

4. ☒ No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:
1-23

Remark on Protest

☐ The additional search fees were accompanied by the applicant's protest.

☐ No protest accompanied the payment of additional search fees.

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US97/11719

BOX II. OBSERVATIONS WHERE UNITY OF INVENTION WAS LACKING
This ISA found multiple inventions as follows:

This application contains the following inventions or groups of inventions which are not so linked as to form a single inventive concept under PCT Rule 13.1. In order for all inventions to be searched, the appropriate additional search fees must be paid.

Group I, claim(s) 1-23, drawn to an immunocontraceptive vaccine composition, comprising sp56 polypeptide, fragments and hybrid molecules thereof, and an immunocontraceptive method.
Group II, claim(s) 25-27, drawn to antibodies against fragments of sp56 polypeptide.
Group III, claim(s) 24 and 28, drawn to a method for determining male fertility.

The inventions listed as Groups I-III do not relate to a single inventive concept under PCT Rule 13.1 because, under PCT Rule 13.2, they lack the same or corresponding special technical features for the following reasons:
An international stage application shall relate to one invention only or to a group of invention so linked as to form a single general inventive concept. If multiple products, processes of manufacture or uses are claimed, the first invention of the category first mentioned in the claims of the application will be considered as the main invention in the claims, see PCT article 17(3) (a) and 1.476 (c), 37 C.F.R. 1.475(d). Group I will be the main invention. After that, all other products and methods will be broken out as separate groups (see 37 CFR 1.475 (d)).
Group I, claims 1-23 form a single inventive concept. Group II is an additional product, an antibody against sp56 polypeptide fragments, and group III is an additional use claimed for said antibodies.